STUDIES ON THE STRUCTURE, LIPID MODIFICATION AND INTERACTIONS OF THE VIRULENCE-ASSOCIATED PROTEINS OF *RHODOCOCCUS EQUI*

TEBEKEME OKOKO

PhD

2014

STUDIES ON THE STRUCTURE, LIPID MODIFICATION AND INTERACTIONS OF THE VIRULENCE-ASSOCIATED PROTEINS OF *RHODOCOCCUS EQUI*

TEBEKEME OKOKO

A thesis submitted in partial fulfilment
of the requirements of the
University of Northumbria at Newcastle
for the degree of Doctor of Philosophy

Research undertaken in the
Faculty of Health and Life Sciences

September, 2014.

ABSTRACT

Rhodococcus equi is a Gram-positive soil organism that causes an aggressive bronchopneumonia in foals and opportunistic infections in immuno-compromised humans. Virulent strains possess an 80 – 90 kb plasmid that encodes an immunogenic surface-located virulence-associated protein VapA. The virulence of the organism has been largely attributed to this protein since mutants lacking *vapA* are attenuated for virulence in mice. VapA is an unusual lipoprotein and existing evidence suggests its biogenesis may not involve normal lipoprotein processing.

In order to understand the structure of VapA and other virulence associated proteins, their genes were cloned, expressed, purified and crystallised. VapG produce high quality crystals that diffracted to 1.8 Å. The structure was resolved to be a closed β-barrel with a long unstructured *N*-terminus which is similar to both VapB and VapD which have also recently been characterised.

Analysis of covalently-bound lipids extracted from membrane proteins that included VapA revealed their modification by mycolic acids, the characteristic branched long-chain fatty acids that define the outer wall of this organism. In order to determine whether a mechanism for mycolylation of surface-exposed lipoproteins might be conserved across the mycolic acid containing actinobacteria, *vapA* was expressed in *Corynebacterium glutamicum*. Here the abundance of protein-bound mycolic acid in VapA-containing fractions increased with *vapA* overexpression. The development of a corynomycoloyl protein transferase assay using surface lipids from various mycolic acid containing bacteria with the characterised enzyme of *C. glutamicum* was attempted. Although the modification of VapA was not observed, an enzyme-lipid adduct seemed to form.

VapA production is coordinated with that of other vaps which are secreted. However the cell bound protein VapA is most critical for virulence. Despite its ascendency, VapA is not sufficient for virulence. It was considered that other Vaps might associate with VapA at the cell surface to exert their effect. Interactions between VapA and the other Vaps were monitored using surface plasmon resonance. VapE, VapG and VapH interacted with VapA in a concentration-dependent manner while VapD seem not to interact with VapA. The implications of the findings are discussed.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgement

I am grateful to a lot of personalities who contributed in no-small measure in the course of my programme.

I would thank my principal supervisor Dr Lynn G. Dover who guided me throughout my research work. I also appreciate the roles of my other supervisors Professor Iain C Sutcliffe and Professor Gary W. Black. Their expertise and dedication were indeed exceptional. I also thank Dr Alistair Brown for providing some technical assistance when consulted. My colleagues have always been wonderful in the laboratory. Their laughter and jokes made the going a little bit easier whenever I got stuck along the way.

I would also thank Vijaya Nataraj of the School of Biosciences, The University of Birmingham for his assistance with analytical ultracentrifugation; Professor Anthony J. Wilkinson, Dr Jean L. Whittingham and Dr Elena V. Blagova of the Structural Biology Laboratory, University of York for their expertise in X-ray Crystallography and structure refinement. I will also appreciate the assistance offered by Dr Helen Waller of Cell and Molecular Sciences, Newcastle University, with surface plasmon resonance, Dr Kaveh Imami of the School of Marine Sciences, Newcastle University with MALDI-TOF analysis, Albert Haas for the provision of *R. equi KasA* mutant, Shinji Takai for monoclonal antibody Mab10G5 and Lothar Eggeling for pEKEx2 plasmid DNA.

I would say a big thank you to my Dad for the encouragement he offered me throughout the programme. My siblings have always been wonderful whenever they were called to duty.

And to my best friend and my wife, Grace who had to cope with my long absence especially with the arrival of Debbie. She has been with me throughout the journey, sharing both the highs and lows all the way. My gratitude to her is incalculable.

My thanks also go to my sponsors, The Tertiary Education Trust Fund (TetFUND) in collaboration with the Niger Delta University for given me this wonderful opportunity of pursuing a research degree.

And to YOU, the Source of Inspiration. I wouldn't be without YOU and if I grow, the reward is YOURS.

Dedication

This work is dedicated to the evergreen memory of my mum Ileimokumo Okoko. She was indeed excited that I was been offered a place for a research degree. Sadly the cold hands of death took her away from me at the peak of the preparations for departure for my programme. But I am thankful to her for imbibing in me the never give-up attitude that defined her.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Name: Tebekeme Okoko

Signature:

LIST OF ABBREVIATIONS

Abbreviations	Meaning
Å	Angstrom
ACN	Acetonitrile
APS	Ammonium persulphate
ATP	Adenosine tri phosphate
AUC	Analytical Ultracentrifugation
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BHI	Brain Heart Infusion
BLAST	Basic local alignment search tool
bp	Base pair(s)
BSA	Bovine serum albumin
CDS	Coding sequences
CFE	Cell free extract
CM	Carboxymethylated dextran
C-terminal	carboxy terminal
Da	Dalton
DMF	Dimethylformamide
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDC	*N*-3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride

EDTA	Ethylene diamine tetraacetic acid, disodium salt
FPLC	Fast protein liquid chomatography
ESI	Electrospray ionization
ExPASY	Expert protein analysis system
g	Gram(s)
h	hour(s)
HBS	HEPES buffered saline
HEPES	*N*-[2-Hydroxylethyl]piperazine-*N*'-[2-ethanesulphonic acid
IEC	Ion exchange chromatography
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IL	Interleukin
INF-γ	Interferon-gamma
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
k_a	Association constant
kb	Kilobase pair(s)
k_d	Dissociation constant
K_D	Equilibrium dissociation constant
kDa	Kilodalton
L	Litre(s)
LB	Luria-Bertani medium
LC-MS/MS	Liquid chromatography Tandem mass spectrometry
Lgt	Prolipoprotein diacylglycerol transferase

Lsp	Lipoprotein signal peptidase
MALDI-TOF	Matrix-assisted laser desorption ionisation time-of-flight
MES	2-(N-morpholino)-ethanesulphonic acid
mg	Milligram
min	Minute(s)
mL	Millilitre
mM	Millimolar
nM	Nanomolar
*M*r	Relative molecular mass
MS	Mass spectrometry
MW	Molecular weight
MWCO	Molecular weight concentrator
m/z	Mass-to-charge ratio
NBT	Nitroblue tetrazolium salt
NCBI	National Centre for Biotechnology Information
ng	Nanogram
NHS	*N*-hydroxysuccinimide
nm	Nanometer
nt	Nucleotide(s)
N-terminal	Amino terminal
OD	Optical density
ORF	Open reading frame
PAGE	Polyacryamide gel electrophoresis
PAI	Pathogenicity Island

PBS	Phosphate buffered saline
PBST	Phosphate buffered saline containing 0.05% Tween-80
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PMSF	Phenylmethanesulfonyl flouride
RMS Δ	Root mean square deviation
rpm	Revolutions per minute
sec	Second(s)
Sec	General Secretory Pathway
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sp.	Species
SPR	Surface plasmon resonance
TAE	Tris-Acetate-EDTA
TAT	Twin-arginine translocation
TEMED	*N',N',N',N'*-tetramethylethylene diamine
TFA	Trifluoroacetic acid
TLC	Thin layer chromatography
Tm	Melting temperature
TNF-α	Tumour necrosis factor alpha
Tris	tris(hydroxymethyl)aminomethane
T7 *lac*	lac operator just downstream of T7 promoter
UV	Ultraviolet

V	Volt(s)
VAP	Virulence associated protein
v/v	Volume per volume
w/v	Weight per volume
WT	Wild type
x g	Times gravity (centrifugal force)
α	Alpha
β	Beta
Δ	Mutant
°C	Degree Celsius
%	Percent
µ	Micro
µL	Micro litre
µg	Micro gram
µM	Micromolar

CHAPTER ONE

1 *Rhodococcus equi*: Virulence, Pathogenesis and Genetics

1.1 Rhodococcus equi: An animal and human pathogen

Rhodococcus equi was first isolated from Swedish foals suffering from bronchopneumonia (Magnusson, 1923). The organism is one of the most important causes of a sub-acute to a chronic suppurative bronchopneumonia, lung abscesses, and several extrapulmonary disorders in foals (Figure 1.1). The host range of the bacterium is relatively broad; it has been isolated from various animals including pigs, sheep, wild boars, cats and dogs (Makrai *et al*., 2002; Takai *et al*., 2003; Ribero *et al*., 2011; Rzewuska *et al*., 2014). However the greatest burden of rhodococcosis (*R. equi* disease) is in foals between the ages of 1 and 6 months, with most of these showing clinical signs of the disease before the age of four months (Nordmann *et al.*, 1992; Giguere and Prescott, 1997). If not treated, mortality rates could approach 80 % (Elissalde *et al.*, 1980; Cohen *et al.*, 2005; Coulson *et al.*, 2010), and it has been reported that foal deaths from pneumonia caused by *R. equi* account for approximately 3 % of foal deaths per annum (Oldfield *et al.*, 2004). The disease is also a major cause of wastage in foals and costs the equine industry millions of dollars to treat (Muscatello, 2012).

Figure 1.1 Right lung from a foal with severe purulent pneumonia caused by *R. equi*.
There is marked consolidation of the cranial portion of the lung and multiple abscesses are present (Vázquez-Boland *et al*., 2013)**.**

Rhodococcus equi is a facultative intracellular organism that is widespread in horse droppings and their environment, thus the organism is isolated primarily from clinical specimens and manure-rich environments (Letek *et al.*, 2010). It tolerates a wide pH range but its growth is optimal between pH 8.5 and 10. This alkaline pH is similar to that of untreated manure and this tolerance potentially provides a selective advantage for colonisation of the farm habitat (Letek *et al.*, 2010). The organism has simple growth requirements that are easily met by organic acids found in herbivore manure and warm temperatures (Barton and Hughes, 1984; Hughes and Sulaiman, 1987; Prescott, 1987).

Two sources for transmission *via* aerosol inoculation are apparent; through inhalation of dust or the breath of companions in farms inhabited by infected horses (Mosser

and Hondalus, 1996; Hondalus, 1997; Letek *et al.*, 2010). However, much remains to be defined; even though the disease is widely reported across horse breeding farms, it remains absent in some farms while in other farms where the disease is reported, only few foals develop the disease (Buntain *et al.*, 2010).

Extrapulmonary disorders (EPDs) that have been reported in horses may be associated with or independent of bronchopneumonia (Reuss *et al.*, 2009). These include diarrhoea, osteomyelitis, hyperthermia and immune-mediated haemolytic anaemia. Interestingly, these may represent the first clinical abnormality observed or they may arise after treatment of the pneumonia. In the latter instance, EPDs may be the ultimate cause of death or prompt for euthanasia (Reuss *et al.*, 2009). Metastatic spread of the organism from the lungs or gastrointestinal tract could be responsible for the reported observed EPDs (Vazquez-Boland *et al.*, 2013).

The administration of rifampin was initially considered to be the treatment of choice for *Rhodococcus equi* infections based on *in vitro* susceptibility data, pharmacokinetic studies and retrospective studies (Giguere *et al*., 2011). Treatment involving a combination of antibiotics (macrolides and rifampin) has proved to be effective against antibiotic-resistant strains which have been isolated from horse-breeding farms (Giguere *et al*. 2012, Burton *et al.*, 2013). However, therapy involving antibiotics with streptolysin-O (as adjuvant) has been shown to be superior to treatment involving antibiotics alone (Horohov *et al*., 2011, Gurel *et al*., 2013).

Though a veterinary pathogen, the organism has emerged as an opportunistic human pathogen as it causes significant morbidity in human patients that are immunocompromised and are exposed to livestock and a dry soil environment (Yamshchikov *et al.*, 2010). Thus HIV-infected patients, individuals receiving organ

transplants and cancer patients are highly susceptible to *R. equi* infections (Prescott, 1991; Perez *et al.*, 2002). In these immunocompromised patients, the clinical manifestations are diverse but about 80 % have pulmonary involvement (Meeuse *et al.*, 2007). However, *Rhodococcus equi* infections have been reported recently in immunocompetent individuals where it manifested as granulomatous mastitis, polymicrobial infections, brain abscesses and abdominal infections (Dias *et al.*, 2013; Nath *et al.*, 2013; Velazquez Benito *et al.*, 2013).

1.2 Morphology and Taxonomy of Rhodococcus equi

Rhodococcus equi is a non-motile, non-spore forming globally-distributed Gram-positive soil organism. It is a coccobacillus especially on solid media (Figure 1.2), but is also pleomorphic with long rods or filaments, rudimentary beading and mycelial branching in liquid media (Yamshchikov *et al.*, 2010). The organism cannot oxidise carbohydrates and alcohols thus cannot utilise sugars as sources of carbon – a property that differentiates it from many *Corynebacterineae* such as *R. jostii*, *R. erythropolis* and *R. opacus* (Letek *et al.*, 2010). It is catalase, urease, lipase and phosphatase positive, but DNAse, elastase, and lecithinase negative (Prescott, 1991; Bidaud *et al.*, 2012). *Rhodococcus equi* grows on non-selective media to a pink– red, irregular,

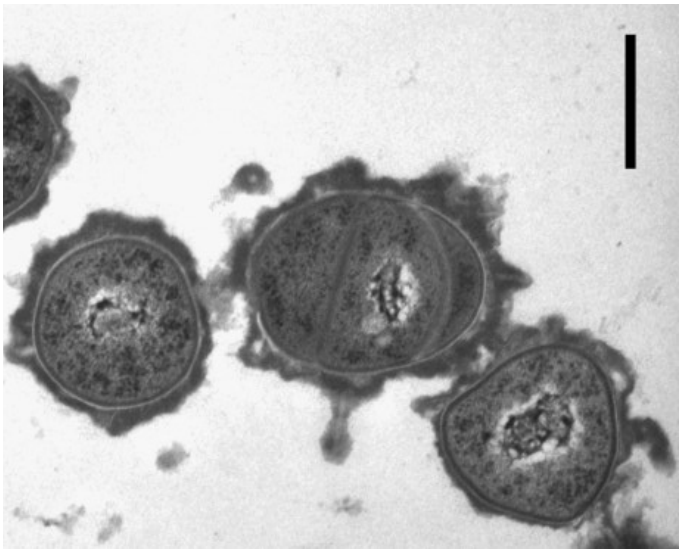


Figure 1.2 Ultrastructural morphology *of Rhodococcus equi* in transmission epectron microscopy

Virulent, plasmid-positive *R. equi* 103+ of a late exponential phase culture were fixed, stained with ruthenium red and prepared for transmission electron microscopy. The image shows a thick layer of polysaccharide capsule material that is surrounding the complex cell wall and stained by the cationic dye (von Bargen and Hass, 2009). Scale bar = 500 nm.

mucoid and round colony which may vary from 2 mm – 4 mm diameter after 48 hours of incubation (Figure 1.3). It is weakly acid-fast in Ziehl-Neelsen stain; though this may depend on the age of the culture, the growth medium and the staining technique (Barton and Hughes, 1980; Kaura and Mutimer, 1987).

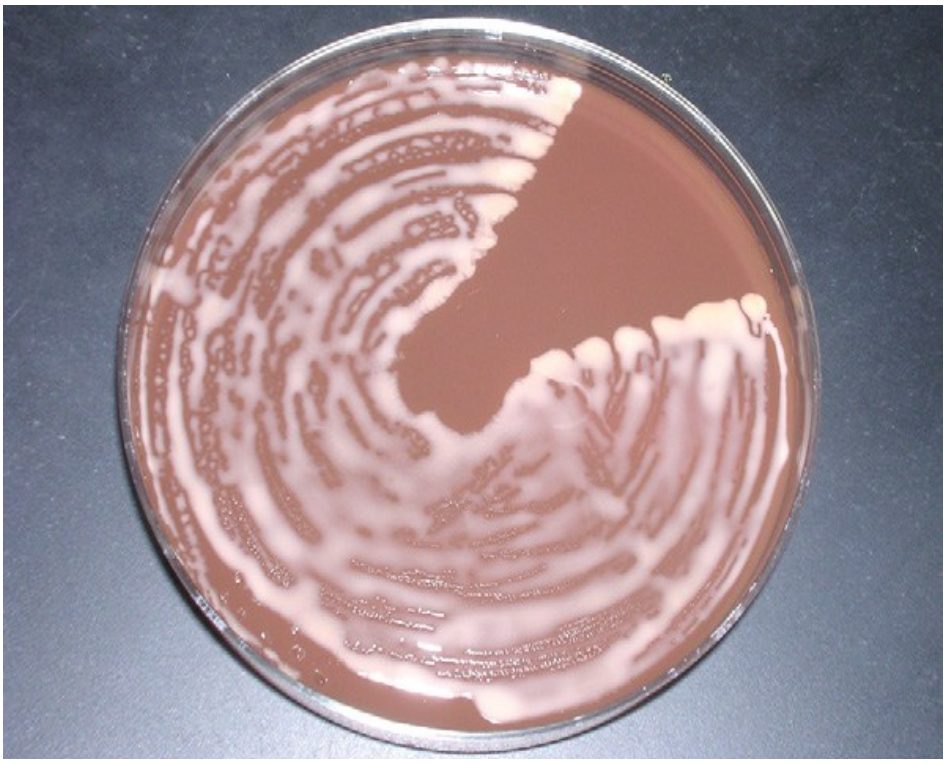


Figure 1.3 Coalescent, mucoid, pink-tinged *Rhodococcus equi* colonies on chocolate agar plate.
From Yamshchikov *et al.* (2010)

The taxon has a confused history. *Rhodococcus* (red-pigmented coccus) belongs to the phylogenetic group referred to as nocardioform actinomycetes that includes *Corynebacterium*, *Nocardia*, and the *Mycobacterium* genera, which are characterised by their high genomic GC content (Goodfellow, 1987; Mosser and Hondalus, 1996).

The name '*Rhodococcus*' was first used by Zopf in 1891 but was redefined in 1977 as a genus that resembles but did not belong to the *Nocardia*, *Corynebacterium*, and *Mycobacterium* genera (Goodfellow and Alderson, 1977). However, Magnusson proposed the name *Corynebacterium equi* for isolates obtained from Swedish foals having pyogranulomatous pneumonia in 1923 (Takai, 1997).

The organisms that were initially grouped under the *Rhodococcus* genus were further divided into two genera (as a result of the heterogeneity of the organisms) namely *Gordona* and *Rhodococcus*. *Gordona* comprised of *Rhodococcus bronchialis, Rhodococcus rubropectinus* and *Rhodococcus terrae*, while all other species were grouped under *Rhodococcus*. However, *Rhodococcus equi* is the only species found in the soil with significant pathogenic potential on horses and other animals including humans (Prescott, 1991; McNeil and Brown, 1994; Mosser and Hondalus, 1996). Recently, *Rhodococcus equi* has been considered to be a new genus with *Prescottella equi* as the proposed name of the type strain. The rationale for this reclassification is that *R. equi* isolated from various sources exhibit minor differences based on phenotypic and genotypic features (Jones *et al.*, 2013 a, b). Furthermore the strains are relatively distantly-related to other rhodococci and *Norcardia* spp (Sangal *et al.*, 2014) (Figure 1.4). The classification of *Rhodococcus equi* presently is:

Phylum: *Actinobacteria*
Class: *Actinobacteria*
Subclass: *Actinobacteridae*
Order: *Actinomycetales*
Suborder: *Corynebacterineae*
Family: *Nocardiaceae*
Genus: *Rhodococcus*
Species: *equi*

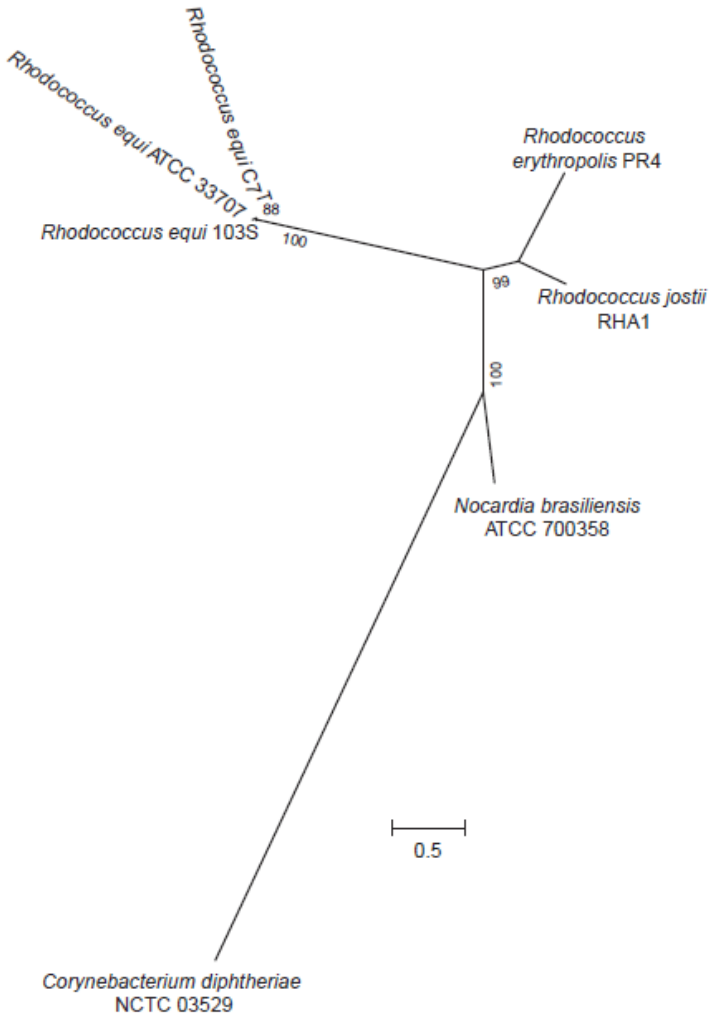


Figure 1.4 Phylogenetic tree from 400 universal proteins showing the genomic relationship of *Rhodococcus equi* to representatives of closely related taxa.
Scale bar shows normalized fraction of total branch lengths (Sangal *et al*., 2014).

1.3 Genome and Virulence Associated Proteins

The genome of *R. equi* 103S consists of a circular chromosome of 5,043,170 bp with 4,525 predicted genes and a circular virulence plasmid of about 80 – 90 kb with an overall GC content of about 68.76 % (Letek *et al.*, 2010) (Figure 1.5). Though the mechanisms of virulence of *R. equi* are not completely understood, there is evidence that it largely depends on the circular virulence plasmid (Meijer and Prescott, 2004)

since strains that are avirulent in horses lack the plasmid (Takai *et al.*, 2000). Sequence analysis shows that this virulence plasmid has 69 open reading frames which span along four regions though only three regions encode proteins with known functions (Figure 1.5). One of the functional regions is involved in plasmid replication and another in conjugation (Meijer and Prescott, 2004: von Bargen and Hass, 2009). The third region of the plasmid (27.5 kb in size) is the pathogenicity island which is characterised by a lower GC content than the rest of the plasmid (Merjer and Prescott, 2004).

This pathogenicity island (PAI) (Figures 1.6) is made up of 26 putative ORFs and contains a gene family that encodes a family of virulence-associated proteins (Vaps) namely VapA, C, D, E, F, G and H (Figure 1.7) (Meijer and Prescott, 2004). Other Vaps namely VapF, pseudo-VapE (VapX) and VapI are frame-shift mutants and thought to be non-functional (von Bargen and Hass, 2009). The Vap genes *vapA*, *vapC*, *vapD*, *vapE*, *vapF*, *vapG* and *vapH* are found over a span of 19,000 base pairs (Meijer and Prescott, 2004). *vapA*, *vapC*, *vapD* are clustered in close proximity to each other while *vapE* and *vapF* are immediately adjacent to each other. However, *vapG* and *vapH* are present as individual genes with *vapG* found in the negative strand (Takai *et al.*, 2000). *vapI*, (ORF13) located just downstream of *vapA*, interrupts the cluster of *vapA*, *vapC*, *vapD* (Polidori and Hass, 2006). Apart from the *vap* locus, other genes which encode proteins with known functions are *orf3* (encoding a SAM-dependent methyl transferase); *lsr2* (encoding a homologue of the mycobacterial nucleoid associated protein); *vcgB* (encoding a protein conserved in pathogenic mycobacteria); *scm2* (encoding a chorismate mutase); and the *virR* operon (Vázquez-Boland *et al.*, 2013). The remainder of the non-*vap* genes encode

products of unknown functions (Ren and Prescott, 2003; Vázquez-Boland *et al.*, 2013).

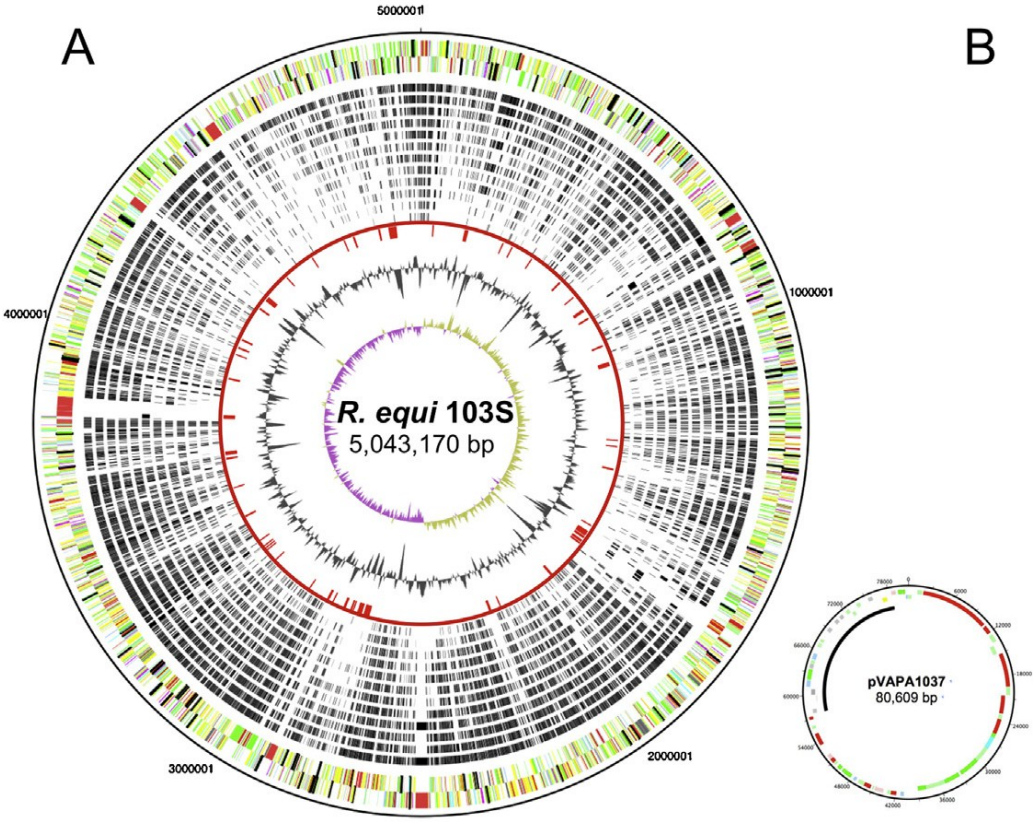


Figure 1.5 The *R. equi* genome.
Circular diagram of (A) *Rhodococcus equi* 103S chromosome and (B) virulence plasmid (not represented at scale). Outer two rings (in chromosome), coding sequences in the forward and reverse strand. Inner rings in the chromosome, orthologue plots from 13 actinobacterial genomes: *R. jostii* RHA1, *Nocardia farcinica* IFM10152, *Mycobacterium smegmatis* MC^2155, *Streptomyces coelicolor* A3(2), *Mycobacterium tuberculosis* H37Rv, *Arthrobacter* sp. FB24, *Corynebacterium glutamicum* ATCC 13032, *Thermobifida fusca* YX, *Frankia sp*. CcI3, *Corynebacterium diphtheriae* NCTC 13129, *Propionibacterium acnes* KPA171202, *Bifidobacterium longum* NCC2705 and *Tropheryma whipplei* TW08 27. Horizontally acquired (HGT) islands are shown in red joined by a red circle; note they tend to coincide with void areas in the orthologue plots, indicating they are not part of the vertically evolving actinobacterial core genome. Inner plots: G + C% (gray) and G + C skew (violet/yellow, origin of replication is clearly detectable). In the plasmid (**B**), the *vap* PAI is represented as a black line (Adapted from Vázquez-Boland *et al.*, 2013).

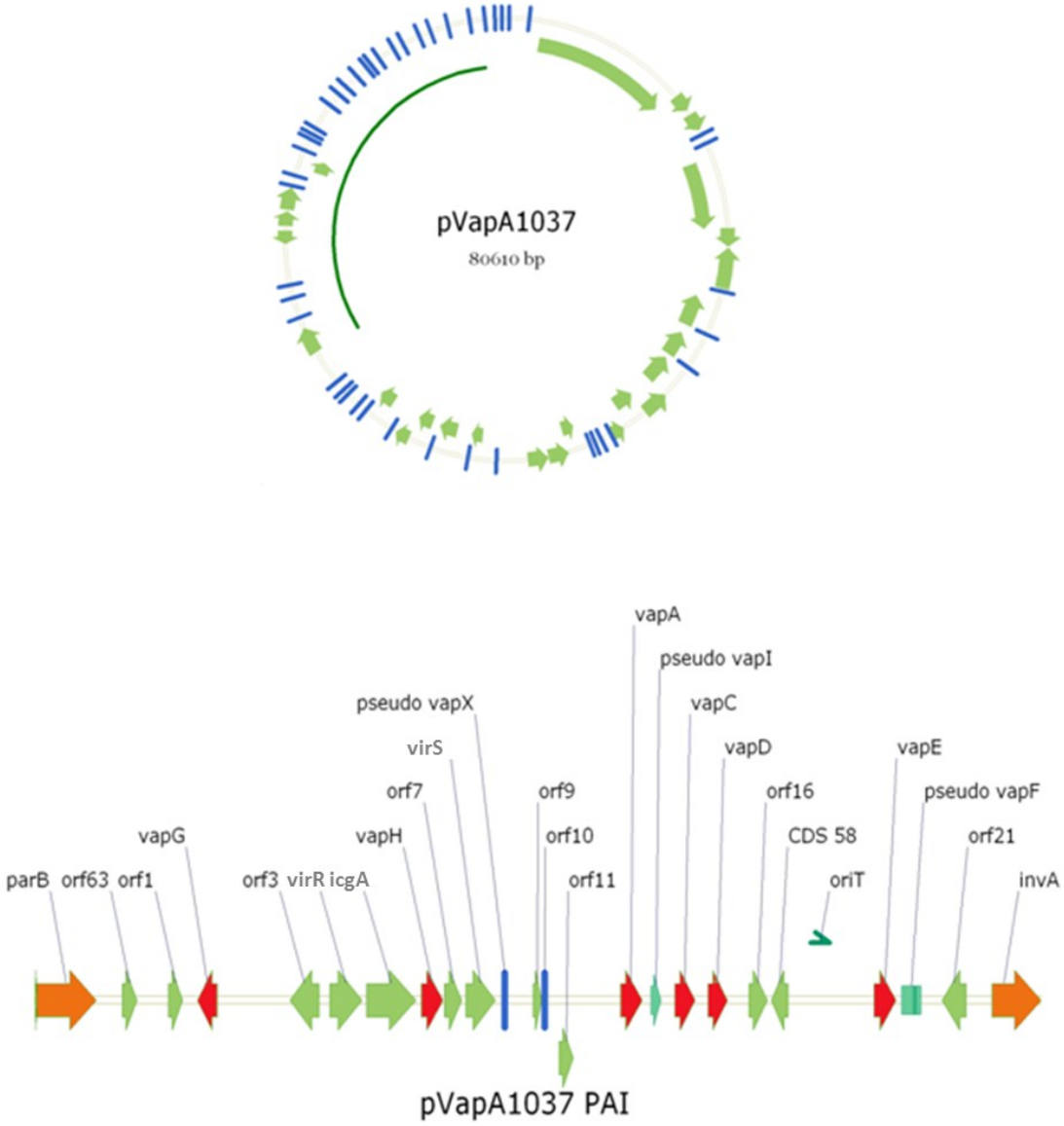


Figure 1.6 Structure of virulence plasmid pVAPA1037 of *R. equi* 103S strain. The circular map shows the position of the pathogenicity island (PAI) (dark green bar) and the positions of the open reading frames. The linear map provides detail of the PAI with genes having known or suspected function named. Functional *vap* genes are coloured in red, the resolvases that limit the PAI in orange, and pseudogenes in turquoise. Very small ORFs or features are illustrated as blue bars.

vapA encodes a protein of 180 amino acids, *vapC* 174, *vapD* 164, *vapE* 189, *vapF* 150, *vapG* 172, *vapH* encodes a protein of 187 amino acids. The virulence of the

organism is largely attributed to the pathogenicity island because VapA (a 15 – 17 kDa antigen) is thought to mediate intracellular growth of the organism in macrophages and disease development *via* the arrest of phagosome maturation. It has been reported that deletion of *vapA* attenuates *R. equi* virulence to almost the same degree as that of the deletion of the entire *vap* locus (Jain *et al.*, 2003). Active immunisation of attenuated *Salmonella* expressing VapA confers long-lasting protection against experimental rhodoccocosis (Cardoso *et al.*, 2013).

Rhodococcus equi isolated from the submaxillary lymph nodes of infected pigs contains one of five plasmids of 79 to 100 kb that encode an 18 - 20 kDa variant of VapA known as VapB (Takai *et al.*, 1996). These isolates exhibit intermediate virulence in mice when compared to the VapA-producing strains. VapA and VapB share 78 % homology (Figure 1.7) and they are allelic variants of one locus that has divergently evolved in two different plasmids. The similarity of VapA and VapB raises the possibility that they could have analogous functions in different strains of *R. equi* (Bryne *et al.*, 2001). They are thus mutually exclusive and also exhibit significant immunological cross reactivity hence the two types of virulence plasmid could have evolved from a single episome (Figure 1.8) (Byrne *et al.*, 2001; Oldfield *et al.*, 2004; Letek *et al.*, 2008). VapB is made up of 197 amino acids. Other *vaps* present in the pathogenicity island of the VapB producing strains are *vapJ* which encodes a protein of 183 amino acids, and *vapK* which encodes a protein of 202 amino acids (Letek *et al.*, 2008).

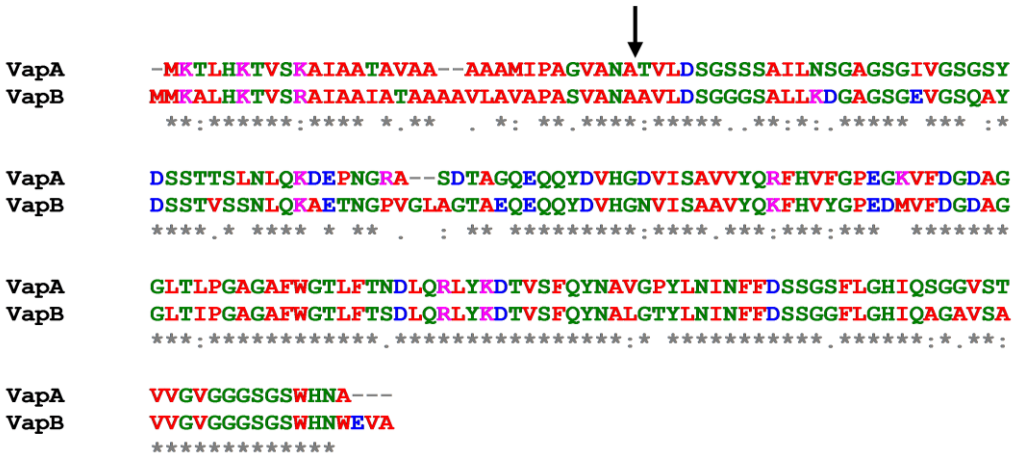


Figure 1.7 Comparison of the amino acid sequences of VapA and VapB.
Sequences were aligned using the Clustal 2.1 sequence alignment program. These share 78 % homology. Arrow shows potential cleavage site for signal peptidase for both proteins.

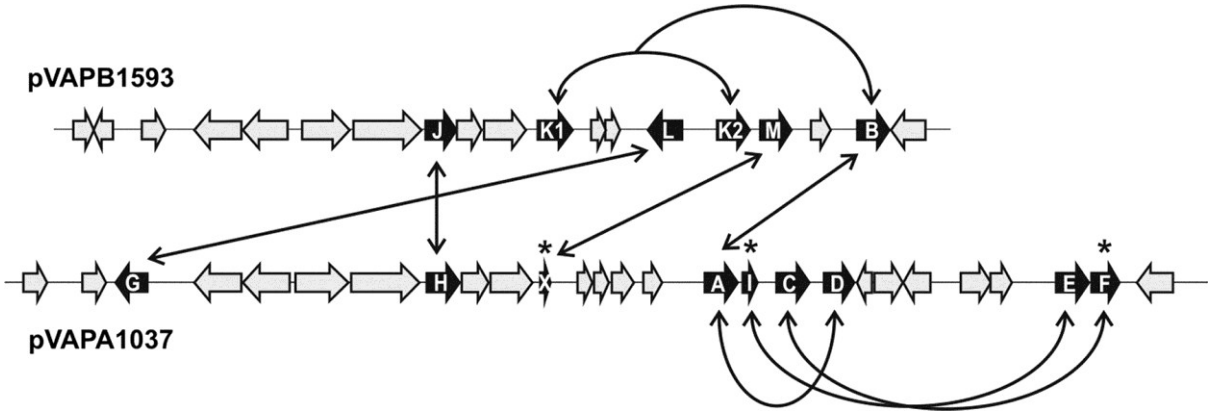


Figure 1.8 Model of the evolutionary dynamics of the *vap* multigene family.
On a schematic representation of the vap PAIs from pVAPB1593 (*vapB* type plasmid) and pVAPA1037 (*vapA* type plasmid), *vap* genes (in black) presumably derived from vertical evolution of a pre-existing common ancestral determinant are connected by straight arrows, and those probably originated by gene duplication in the specific plasmid are connected by curved arrows (see the text for details). The connection between *vapM* and the *vapX* pseudogene was deduced from a phylogenetic tree constructed with the 28-residue VapX product aligned with the corresponding protein fragments from other members of the Vap family (not shown). Asterisks indicate degenerate *vap* genes; note in pVAPA1037 that one gene in each of the duplicated *vapIE* and *vapCF* tandems is undergoing decay; pVAPB1593 is the *vapB* type plasmid; pVAPA1037 is the *vapA* type plasmid (Adapted from Letek *et al.*, 2008).

1.4 Pathogenesis and pathology of Rhodococcus equi infections

A key virulence determinant is the possession of the virulence plasmid described in detail in Section 1.3. Several strains of *Rhodococcus equi* have been identified based on plasmid diversity but the organism can be grouped into *vapA*-expressing strains, *vapB*-expressing strains and those that neither express *vapA* nor *vapB i.e.* that possess no virulence plasmid (Meijer and Prescott, 2004; Poolkhet *et al.*, 2010).

The *vapA*-containing strains cause rhodococcosis in young horses where the most common manifestation is a chronic suppurative bronchopneumonia with extensive abscessation (Figure 1.2) and associated suppurative lymphadenitis (Phumoonna *et al.*, 2006). *In vitro* experiments suggest that the infectivity of *Rhodococcus equi* is limited to the monocyte-macrophage lineage (Mosser and Hondalus, 1996) thus alveolar macrophages could be the main targets of the organism. The early stages of the infection resemble those of other lung infections thus may pose difficulties in early diagnosis coupled with the slow spread of the lung infection (the organism being a slow grower itself) (Giguere and Prescott, 1997). In fact, *R. equi* infections have been misdiagnosed as mycobacterial and diphtherial infections (Salifu *et al.*, 2013). Early lesions are characterised by cellular influx into the alveolar spaces (which normally consists of macrophages and giant cells with fewer neutrophils) where the bacteria inhabit the phagocytic cells (Hondalus, 1997). Their ability to replicate in macrophages may result in necrosis followed by the destruction of the lung parenchyma at the advanced stages of the disease (Hondalus, 1997). This may result from the destruction of the cell *via* lysosomal degranulation after intracellular multiplication of the organism (Yager, 1987). Though the exact mechanism underlying the survival of the organism is unknown, inadequate acidification of the

phagosome may be responsible for the inability of the phagocytes to degrade the bacteria (Zink *et al.*, 1987; Toyooka *et al.*, 2005). Sometimes, the infection manifests as intestinal pneumonia and, in such cases, the clinical signs could be fever, anorexia, diarrhoea and weight losses which may be severe thus difficult to treat (Zink *et al.*, 1986; Baldwin *et al.*, 1992). The *vapB*-containing *R. equi* mostly cause disease in pigs and cattle which manifests as chronic pyogranulomatous adenitis (de Vagas *et al.*, 2013). These strains also cause human rhodococcal lung infection which resembles pulmonary tuberculosis and is thought to have a high mortality (Poolkhet *et al.*, 2010; Vázquez-Boland *et al.*, 2010). In general, the *vapB*-bearing strains are less virulent than the *vapA*-containing strains (Meijer and Prescott, 2004).

1.5 Rhodococcus equi cell envelope

The basis of the pathogenicity of *R. equi* relies on its ability to multiply inside the macrophage where it eventually causes degranulation of the lysosome and subsequently leads to cell death (Prescott, 1991). The organism has a complex cell envelope which enables it to survive under very harsh conditions such as low pH and oxidative stress; the latter being a major mechanism used by macrophages to combat foreign bodies (Benoit *et al.*, 2000; Benoit *et al.*, 2002).

The rhodococcal cell envelope is dominated by mycolic acids. These branched long-chain fatty acids are covalently linked to a peptidoglycan-arabinogalactan complex and thus form the basis for an anchored outer permeability barrier. Other components not covalently attached to the core are glycolipids, polysaccharides, lipoglycans, lipoproteins and channel-forming porins (Sutcliffe *et al.*, 2010). A model of the rhodococcal cell envelope is shown in figure 1.9. This is heavily based upon the

Minnikin model of the cell wall of mycobacteria (Minnikin, 1982) and developments thereof.

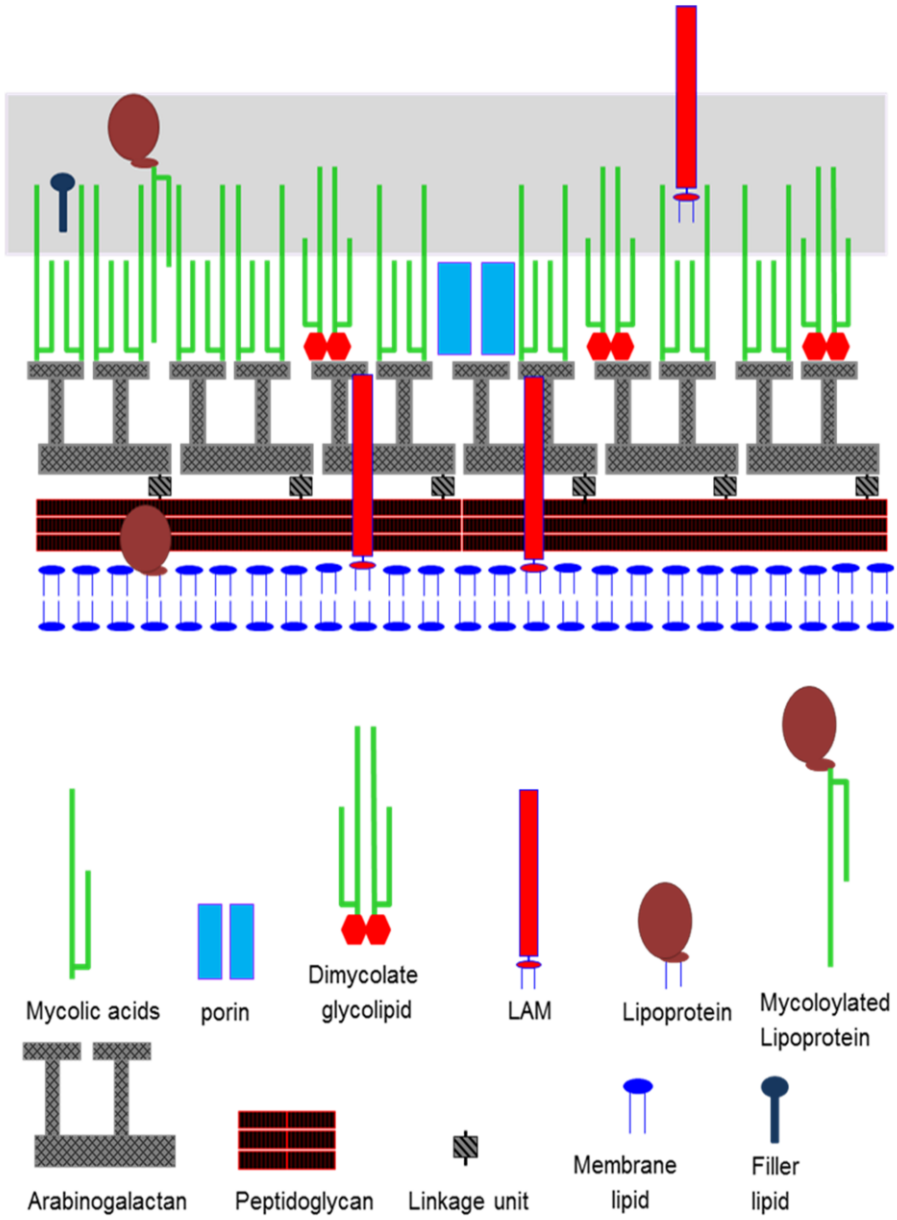


Figure 1.9 Model for the organisation of the rhodococcal cell envelope.
This figure represents an adaptation of the published model by Sutcliffe *et al.* (2010). It illustrates that mycolic acids form the basis of an outer lipid permeability barrier. Individual envelope components are identified beneath the model. Not shown are the outermost layers composed of the cell envelope polysaccharides and/or capsules that are known to be present in many rhodococci. The model emphasises the possibility that lipoproteins which occupy the outer permeability barrier might be mycoloylated.

1.5.1 Mycolic acids

Mycolic acids are high molecular weight α-alkyl, β-hydroxy fatty acids. They vary in size and complexity and are found exclusively in the cell envelopes of different genera of the mycolata (all mycolic acid-containing Actinobacteria) (Sutcliffe, 1998). In *Corynebacterium*, mycolic acids can contain from 22-36 carbon atoms; 34-38 carbon atoms in *Dietzia*; 46-66 in *Nocardia* while those of *Mycobacterium* range between 60 and 90 carbon atoms; *Rhodococcus* can produce mycolic acids in the range of 28-54 carbon atoms (Sutcliffe *et al.*, 2010; Marrakchi *et al.*, 2014). In *R. equi*, the mycolic acids are covalently-bonded to the cell wall arabinogalactan and are also esterified to carbohydrate groups to form glycolipids (Sutcliffe, 1998; Hsu *et al.*, 2011). For the most part however, they are esterified to the terminal arabinose units of the peptidoglycan-linked polysaccharide complex, *i.e.* arabinogalactan (Lee *et al.*, 1997). These long-chain lipids are the primary mediators of the hydrophobic character of the cell envelope and form a closely packed, impermeable monolayer (Tahlan *et al.*, 2012). The mycolic acids of *R. equi* have no functional groups in their meromycolate chains other than the possibility of one or more double bonds (Verschoor *et al.*, 2012). The structures of some representative mycolic acids within the mycolata are shown in figure 1.10.

The biosynthesis of mycolic acids has been a subject of intense investigation but considering their exclusive presence in organisms grouped in the *Corynebacterineae* suborder, the enzymatic steps involved in their biosynthesis may only be common to this group of organisms (Portevin *et al.*, 2004). The reactions that constitute mycolic acid biosynthesis (also known as the mycolic acid pathway) consist of over 200

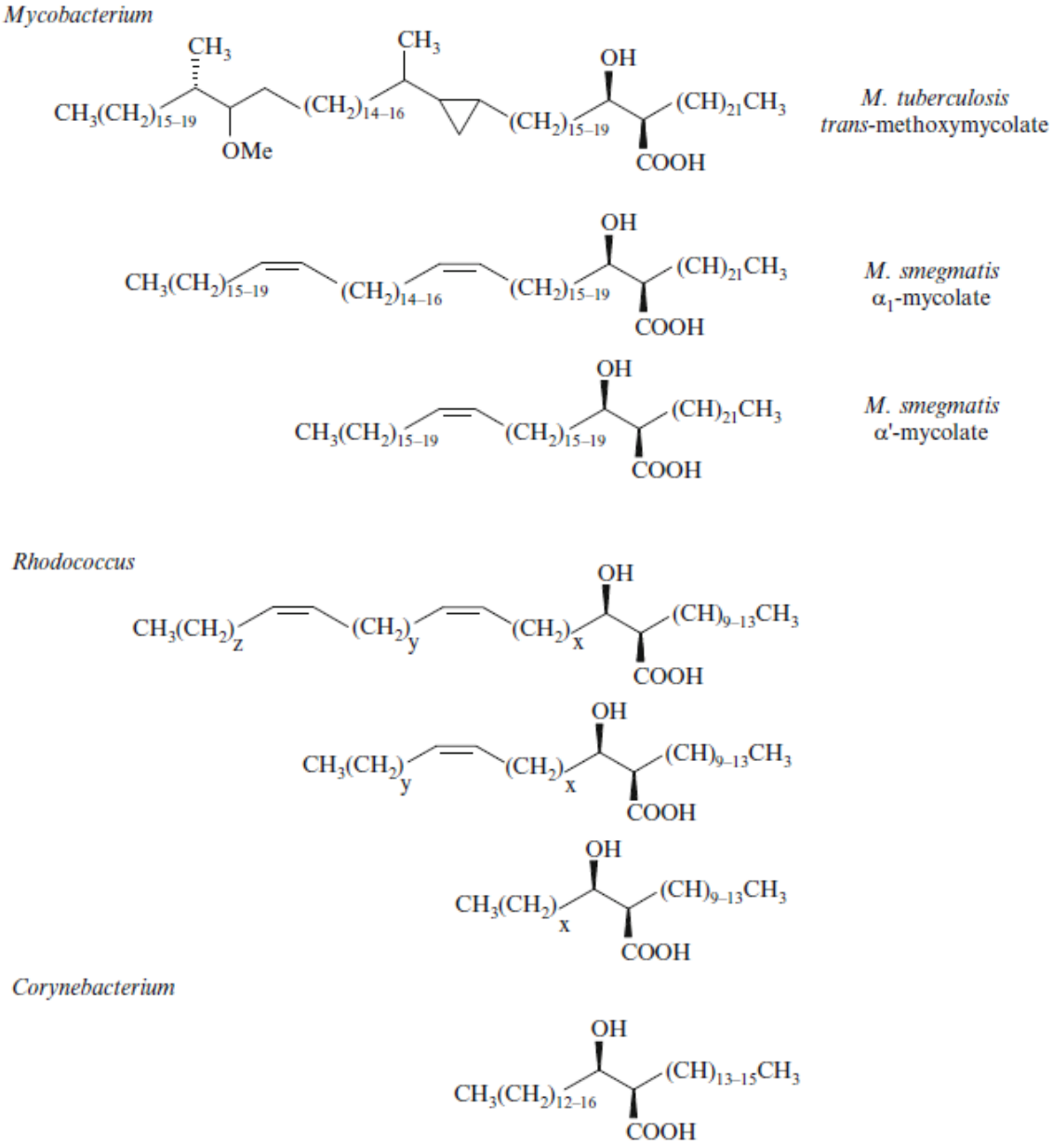


Figure 1.10 Structures of representative mycolic acids from *Corynebacterium*, *Rhodococcus* and *Mycobacterium* species.

Structures illustrate their differing complexity. *Mycobacterium tuberculosis* methoxymycolate is an example of the most complex mycolic acids. The less complex *M. smegmatis* mycolates encompass either double or single unsaturations whilst retaining the longer chain length. *Rhodococcus* sp. mycolates are of an intermediate size. They present an aliphatic 2-alkyl chain varying from 12 to 16 carbons, whilst the 3- hydroxyl meromycolate typically contains 18–40 carbons (*i.e.* x, y and z total 18–40) (Sutcliffe *et al.*, 2010).

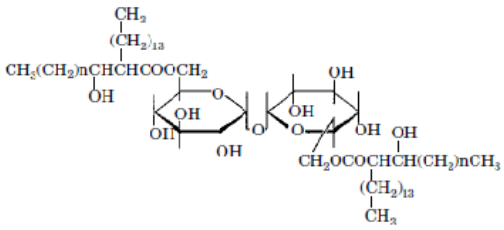
different reactions and up to 200 metabolites mediated through a sequence of proteins (Raman *et al.*, 2005). The reactions of the mycolic acid biosynthesis are grouped under the following major pathways (a) production of malonyl-CoA, (b) fatty acid synthase-I (FAS-I pathway), (c) fatty acid synthase-II (FAS-II pathway) (d) the condensation of FAS-I and FAS-II products (Raman *et al.*, 2005). Acetyl-CoA (produced from other pathways) is carboxylated to malonyl-CoA in a reaction catalysed by acetyl-CoA carboxylase. FAS-I then catalyses the elongation of malonyl-CoA to short-chain acyl-CoAs such as C_{16}, C_{20}, or C_{24} CoAs that could serve as the α-alkyl branch (R_2) (Gao *et al.*, 2003; Bhatt *et al.*, 2007; Grzegorzewicz *et al.*, 2012). For the meromycolate backbone (R_1), products of the FAS-I are further elongated in reactions catalysed by FAS-II. This elongation likely ceases when the chain attains the length required for the meromycolate chain modification or condensation (Marrakchi *et al.*, 2014). In mycobacteria, this chain is further functionalised with *cis* or *trans* double bonds, *cis* or *trans* cyclopropane rings and/or polar functional groups such as methyl ethers, ketones, esters or epoxides (Grzegorzewicz *et al.*, 2012). The elongation reactions are catalysed by a series of β-ketoacyl-ACP synthases which are components of the fatty acyl synthase pathways. The R_1 and R_2 chains are finally condensed in a series of reactions catalysed by polyketide synthase (a subunit of the FAS-II complex) to yield the mycolic acids (Portevin *et al.*, 2004).

In addition to being a probable virulence factor of the organism, mycolic acids also contribute to resistance to antibiotics and offer a considerable hydrophobic cell surface property. This feature is a major determinant of the organism's physiology (Bendinger *et al.*, 1993; Jarlier and Nikaido, 1994; Lee *et al.*, 1997).

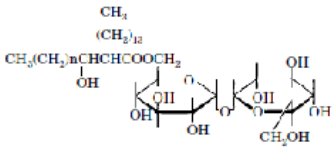
1.5.2 Lipoarabinomannans and Glycolipids

Lipoarabinomannans (LAMs) are lipoglycans that are almost ubiquitous in mycolic acid-containing Actinobacteria. They are amphipathic molecules presenting a tripartite structure which is made up of a mannosyl-phosphatidyl-*myo*-inositol (MPI) lipid anchor that embeds in the outer leaflet of the plasma membrane, a polysaccharide backbone composed of D-mannan and D-arabinose (or D-arabinan) and terminal oligosaccharide caps (Chatterjee and Khoo, 1998; Vercellone *et al.*, 1998). LAM and its related macroamphiphile precursors, lipomannan (LM) and phosphatidyl-*myo*-inositol mannosides (PIMs), are often found together in the cell walls of the mycolata (Besra and Brennan, 1997; Nigou *et al.*, 2003).

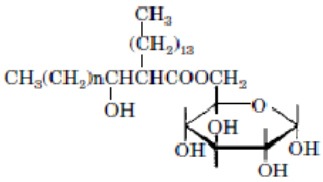
In addition to the LAMs, many glycolipids are bound to the cell walls of members of the mycolata (Sutcliffe *et al.*, 2010). In *Mycobacterium tuberculosis*, some important surface-exposed antigenic glycolipids have been reported. They include the phenolic glycolipids, diacyltrehaloses, and lipooligosaccharides (Munoz *et al.*, 1997). Some of the glycolipids that are surface exposed in the cell wall contain mycolyl residues and thus are termed mycolyl glycolipids (Figure 1.11). Trehalose 6, 6'-dimycolate (TDM), which was first isolated as a 'cord factor' from highly virulent *Mycobacterium tuberculosis*, has been recognised as one of the virulence factors inhibiting the formation of the phagolysosome in infected macrophages (Kai *et al.*, 2007). TDM is also found in many *Actinomycetales* including *Rhodococcus*, *Nocardia* and *Corynebacterium*. Other glycolipids isolated from members of *Rhodococcus* include trehalose 6-monomycolate, glucose 6-monomycolate, mannose 6-monomycolate and fructose 6-monomycolate (Ueda *et al.*, 2001).



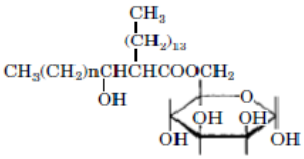
(a) Trehalose 6, 6 'dimycolate (TDM)



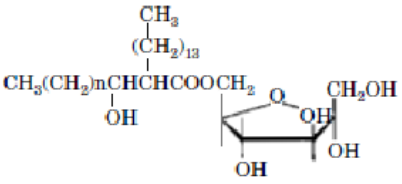
(b) Trehalose 6-monomycolate (TMM)



(c) Glucose 6-monomycolate (GMM)



(d) Mannose 6-monomycolate



(e) Fructose 6-monomycolate (FMM)

Figure 1.11 Glycolipids derived from *Rhodococcus*.
The number of carbon atoms, *i.e.* n could be 16, 18 or 20. (Modified from Ueda *et al.*, 2001).

1.5.3 Channel-forming porins

In *Rhodococcus* and other related mycolata, the presence of the asymmetric outer mycolic acid layer serves as a permeability barrier surrounding the peptidoglycan layer (Riess and Benz, 2000). Therefore this group of organisms must contain features in their cell wall that could allow hydrophilic substances to move in/out of what is termed the pseudoperiplasm (Sutcliffe *et al.*, 2010). These so-called porins produce large and water-filled channels that are permeable to small hydrophilic substances and waste products but are not permeable to larger and more lipophilic compounds such as antibiotics and inhibitors (Nikaido, 1994). The presence of cation-selective pore-forming proteins has been reported in many mycolata such as *Corynebacterium glutamicum*, *Mycobacterium bovis, M. tuberculosis* and *Rhodococcus erythropolis* (Lichtinger *et al.*, 2000). *Mycobacterium bovis* also contains an anion-selective channel in the cell wall (Lichtinger *et al.*, 1999), thus the mycolata contain porins of different specificities. Two porins in the cell wall of *R. equi* have been described by Riess *et al.* (2003). These porins are organic solvent-extractable proteins with complementary selectivity for anions and cations. The cation-selective channel denoted $PorA_{Req}$ is a voltage-gated channel with an apparent molecular weight of 67 kDa and a pore diameter of 1.8 - 2.0 nm while the anion selective channel $PorB_{Req}$ is not voltage-gated and also has a pore diameter of about 2.0 nm but this protein has an apparent molecular weight of only 11 kDa; clearly there are fundamental differences regarding the protein architecture of these channels (Riess *et al.*, 2003). The properties of these pore-forming proteins are similar to those found in the cell walls of other closely-related mycolic acid containing Actinomyces (as described by Riess *et al.* 2003) as shown in Table 1.1.

Table 1.1 Comparison of the cell wall channel properties of *R. equi* and other related Mycolata
(Modified from Riess *et al*., 2003. *From Mafakhari *et al*., 2014)

Species	Porin	G(nS)[1]	Selectivity[2]	Point charges at channel mouth	Channel diameter (nm)
Rhodococcus equi	$PorA_{Req}$	4.00	9.0	-1.5	2.0
	$PorB_{Req}$	0.30	0.16	+1.5	1.4
Rhodococcus erythropolis	PorA	6.00	11.80	-2.7	2.0
Rhodococcus corynebacteroides	PorA	5.50	3.80	-2.7	2.2
Corynebacterium glutamicum	PorA	5.50	8.10	-2.0	2.2
Nocardia farcinica	NfpA	3.0	8.2	-1.3	1.4
*Dietzia maris**	PorA	5.75	20.7	-2.4	2.1

[1]G, Single channel conductance (in nanoSiemens, nS) in 1 M KCl. [2] P_{cation}/P_{anion} measured in KCl.

In *C. glutamicum*, the possession of the cation-selective channel PorA (due to the presence of negatively-charged groups localised at the channel mouth) is solely responsible for the cell wall permeability to hydrophilic substances (Lichtinger *et al.*, 1998). Since *R. equi* is a relative of *C. glutamicum*, $PorA_{Req}$ and $PorB_{Req}$ may also perform roles similar to those of *C. glutamicum*.

1.6 Other virulence factors

1.6.1 Plasmid-encoded factors

The virulence of *R. equi* is largely attributed to *vapA* and its transcription is predominantly under the regulation of proteins encoded by the genes *virR* and *virS*, which are part of the *virR* operon (Byrne *et al*., 2007) (Figure 1.6). Thus the expression of *vapA* is significantly reduced and virulence of *R. equi* is attenuated if there is mutation of either *virR* or *virS* (Giguere *et al.*, 2011; MacArthur *et al.*, 2011).

Another gene, *icgA* is also a component of the *virR* operon (Figure 1.6) that encodes a major facilitator superfamily transport protein. Though deletion of *icgA* did not affect the transcription of *vapA*, it elicited an enhanced growth phenotype in macrophages and a significant reduction of the viability of the organism in macrophages (Wang *et al*., 2014).

1.6.2 Chromosome-encoded factors

Rhodococcus equi chromosome encodes two important enzymes of the aromatic amino acid biosynthesis pathway; chorismate mutase and anthranilate synthase (REQ23860 and REQ23850 respectively). The encoding genes are strongly co-induced with the *vap* genes (Letek *et al.*, 2010). The amino acids tyrosine, phenylalanine and tryptophan appear to be present at limited concentrations inside the hypoxic macrophage intracellular environment, thus *R. equi* relies on *de novo* supply of the aromatic acids. Thus full proliferation and intracellular fitness of the bacterium could be compromised following independent mutations of REQ23860 and REQ23850 (Letek *et al.*, 2010).

Other factors are important in the modification and exploitation of host lipids. The exoenzymes cholesterol oxidase and phospholipase C could also be important virulence factors (Hondalus, 1997). The 54 kDa secreted protein, cholesterol oxidase catalyses the oxidation of cholesterol to 4-cholesten-3-one and is thought to be responsible for the destabilisation of the host cell membrane thus may play a role in tissue damage associated with *R. equi* infections (Linder and Bernheimer, 1997). There are four phospholipase C proteins encoded in the genome of *R. equi* (Vera-Cabrera *et al.*, 2013). The activity of cholesterol oxidase is enhanced by phospholipase C and provides access to the target thus the combined action of

cholesterol oxidase and phospholipase C may confer membranolytic activity of *R. equi*. These two enzymes have been termed 'equi factors' (Prescott, 1991). The target of cholesterol oxidase may be the sterol layer of the lysosomal or cellular membranes of the macrophage and thus it may contribute to the macrophage degeneration observed *in vitro* on challenge with *R. equi* (Hondalus, 1997). *Rhodococcus equi* catabolises cholesterol to products used as substrates for other important biochemical pathways. Mutations in some key enzymes of this pathway of cholesterol catabolism (also called the steroid catabolic pathway) in *R. equi* displayed attenuated phenotypes and reduced proliferation in an *in vitro* macrophage infection assay (van der Geize *et al.*, 2011). *Rhodococcus equi* may use this pathway to utilise macrophage membrane sterols possibly due to the extra demand for carbon in the hostile environment. In addition to cholesterol oxidase and phospholipase, *Rhodococcus equi* contains an abundant array of other secreted lipases which can degrade host cells for fatty acid oxidation (Alvarez *et al.*, 2013; Villalba *et al.*, 2013) (Figure 1.12).

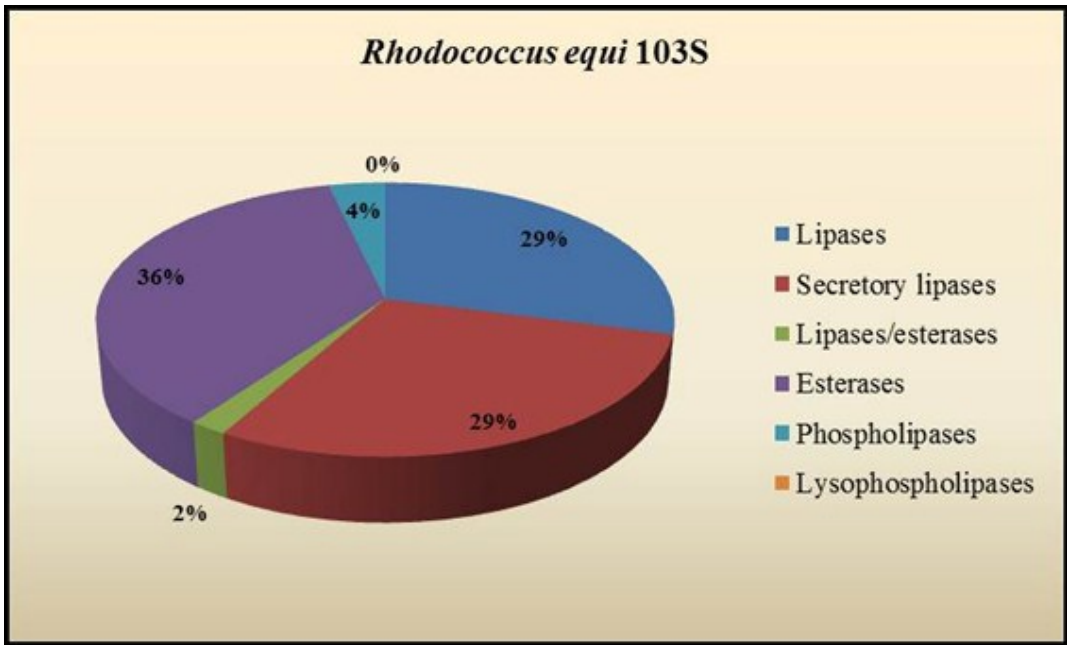


Figure 1.12 Relative proportions of diverse lipolytic enzymes occurring in *Rhodococcus equi*
(Adapted from Villalba *et al*., 2013)

The nitrate reductase encoded by *narG* for could also be a possible virulence factor. Though being an aerobe, *R. equi* is also facultative since it is suited for anaerobic respiration *via* denitrification reducing nitrates to nitrites (Letek *et al.*, 2010). This may suggest that the organism could use nitrate as a terminal electron acceptor in its respiratory chain in the hypoxic macrophage vacuolar compartment. Mutation in *narG* was also found to be significantly attenuated in a mouse challenge study (Giguere *et al.*, 2011).

1.6.3 Environmental factors

Temperature and pH are important determinants of the virulence of *Rhodococcus equi*. The expression of the *vap* PAI genes is transcriptional regulated by temperature

and pH with strong upregulation at 37°C pH 6.5 and downregulation at 30°C pH 8.0 (Vázquez-Boland *et al*., 2013). Other environmental stresses that regulate the expression of genes within the pathogenicity island are oxidative stress, magnesium and iron (Meijer and Prescott, 2004). The ability of the organism to acquire ferric iron from the host is directly related to virulence (Letek *et al*., 2010).

1.7 Immunological aspects of Rhodococcus equi infections

The ability to replicate inside the macrophage is a major determinant of the virulence of *R. equi* (Hondalus and Mosser, 1994). Macrophages may be the prime hosts of the organism since *in vitro* experiments show that the virulence of the organism is limited to cells that express the complement receptor Mac-1 (CD11b and CD18). These receptors are exclusive to the macrophage superfamily (Hondalus *et al.*, 1993). The intracellular growth of the organism has been considered toxic because it inhibits the formation of the phagolysosome either as a consequence of intracellular growth or as a mechanism for survival (Hondalus and Mosser, 1994). Usually, a phagosome is created when a phagocyte engulfs a non-self substance including a bacterium *via* receptor-mediated endocytosis. This phagosome then matures to a phagolysosome *via* an early phagosome and a late phagosome stage by interacting with the lysosome (von Bargen and Haas, 2009). Destruction of the foreign body is then achieved by the combination of low pH (4.0 – 5.0), hydrolytic enzymes, and the production of reactive oxygen and nitrogen species (von Bargen and Haas, 2009). According to Fernandez-Mora *et al.* (2005), the majority (*ca.* 90 %) of the vacuoles containing *R. equi* are arrested at a late endosome stage *en route* to the lysosome thus preventing the maturation of the phagosome (Figure 1.13).

However, Toyooka *et al*. (2005) argued that when virulent *R. equi* are internalised in the macrophage, maturation of the phagosome proceeds initially just like other phagosomes containing harmless microorganisms, but they are tolerant to the bactericidal substances in the phagolysosomal environment. *Rhodococcus equi* does so by secreting substances that hinder the acidification of the phagolysosome thus the optimum pH required for the activity of the phagosolysosome is not attained (Toyooka *et al.*, 2005). Hence the organism continues to multiply in the compartment and eventually causes cell necrosis (Figure 1.14).

When macrophages are activated (maybe by engulfed microorganisms), cytokines, chemokines and other pro-inflammatory mediators are expressed and this could lead to generation of an oxidative burst - a process whereby reactive oxygen and nitrogen species (ROS and RNS) are released, which destroy the microorganisms. Lysozyme, a cell wall degrading enzyme, is an important component of the innate immune system. It has been reported that *R. equi* is moderately resistant to lysozyme (Hebert *et al.*, 2014). This is to be expected as the enzyme would not readily access its peptidoglycan substrate which lies beneath the mycolic acid layer.

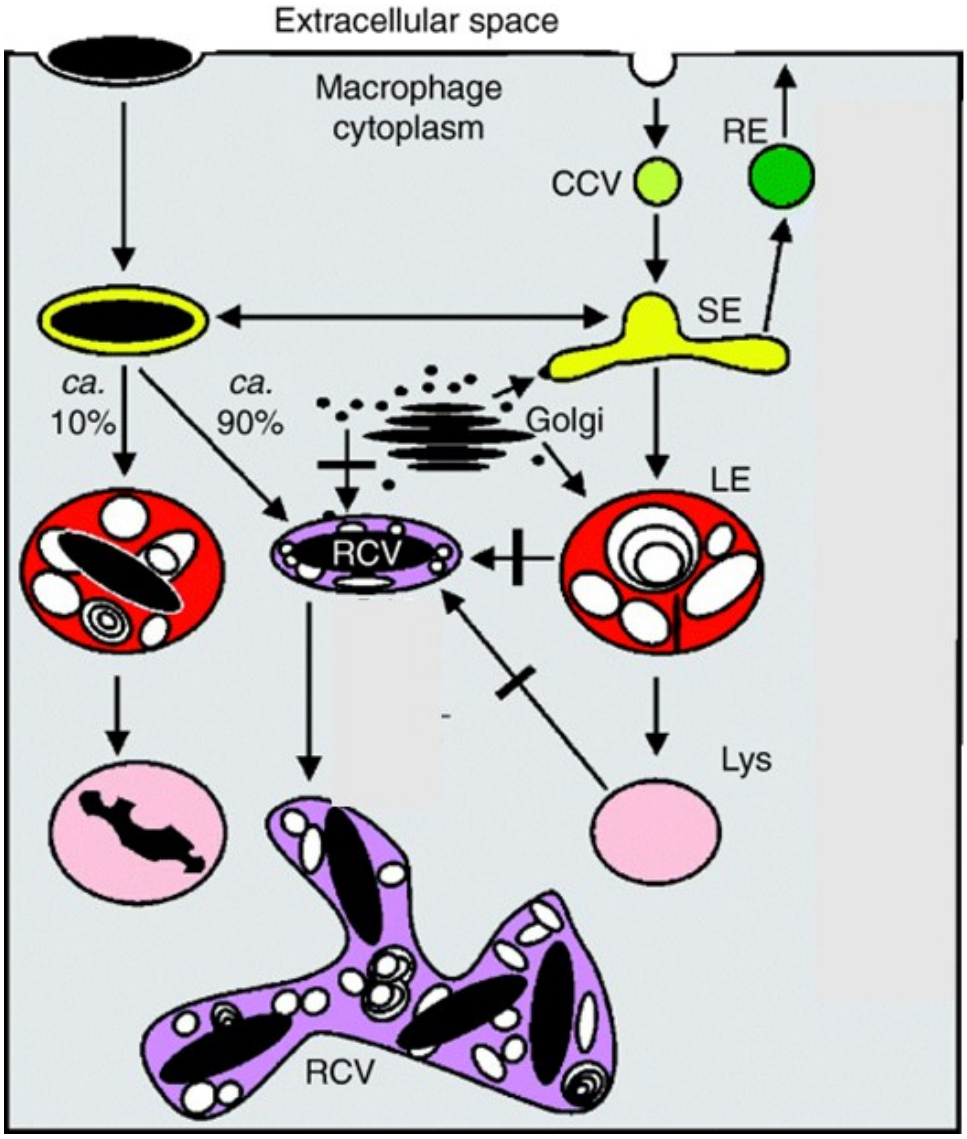


Figure 1.13 A model for *R. equi* infectivity in macrophages.
Clathrin-coated vesicles (CCV) feed extracellular liquid into the endocytic system. Phagosomes are formed at the plasma membrane and normally interact serially with early sorting endosomes (SE), late endosomes (LE) and lysosomes (Lys), but not with early recycling endosomes (RE) through which plasma membrane receptors can be recycled from the SE back to the surface. Some 10% of the ingested *R. equi* are transported along the degradative pathway to a (phago)lysosomal compartment and likely degraded, whereas the majority (some 90%) are directed to an unusual compartment whose composition is between the SE and the LE. The Golgi compartment communicates through vesicular trafficking with the SE/LE system but not with RCVs. Typical compartmental markers are indicated (Modified from Fernandez-Mora *et al*., 2005).

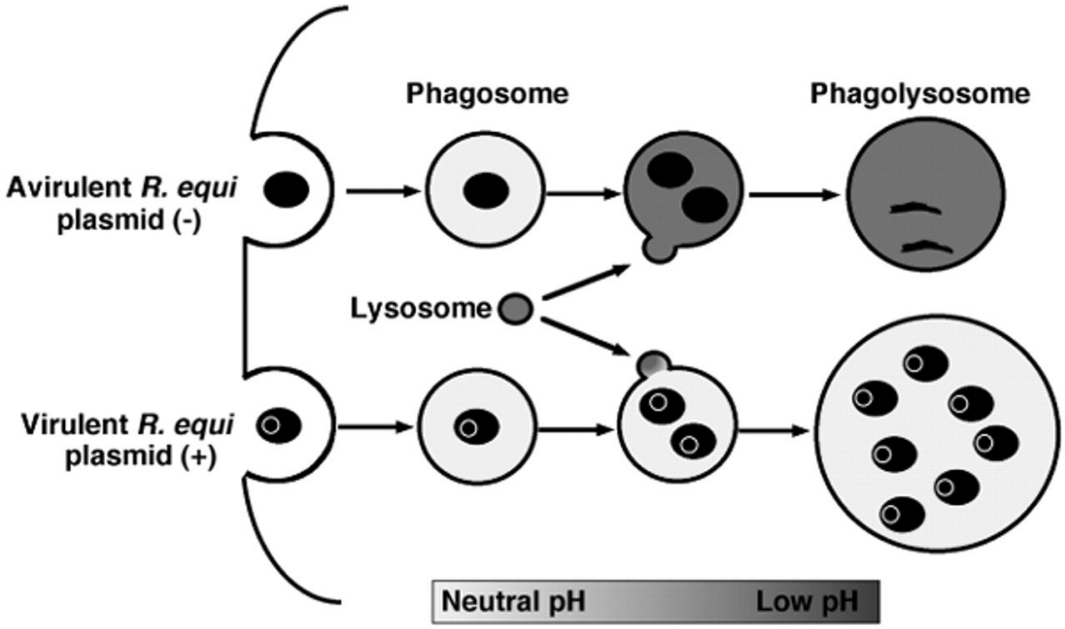


Figure 1.14 Alternative model for *R. equi* infection in macrophages.
After engulfing *R. equi*, phagosomes containing virulent or avirulent *R. equi* fuse with lysosomes. Thereafter, avirulent *R. equi* lacking a plasmid cannot survive an acidic compartment in the phagolysosomes. Virulent *R. equi* harbouring an 85 kb plasmid suppresses the acidification. Thus they survive and replicate in the phagolysosomes (Adapted from Toyooka *et al*., 2005)

Nitric oxide (NO) is an important reactive nitrogen species that is synthesised from arginine in a reaction catalysed by inducible nitric oxide synthase (iNOS). Interferon gamma (IFN-γ) is a primary signal for the transcription of iNOS though there are other secondary signals such as tumour necrosis factor alpha (Darrah *et al.*, 2000). When *R. equi* strains enter the alveolar macrophage of susceptible foals, there is an initial decrease in the transcription of IFN-γ mRNA (Hines *et al.*, 2003). This suggests that *R. equi* may produce factors that suppress the production of IFN-γ. Interferon gamma activates macrophages and upregulates phagolysosomal fusion (as part of the Th1 response) thus the cytokine plays a central role in host susceptibility to *R. equi* with regards to phagocytosis (Kanaly *et al.*, 1995).

Experiments in mouse models involving adoptive transfer and neutralising antibodies have shown that challenge with *R. equi* results in the development of Th2 responses, characterised by production of interleukin-4 (IL-4) rather than IFN-γ. Such mice develop characteristic pulmonary lesions probably because of the dearth of IFN-γ (Patton *et al.*, 2005). Macrophages could be activated to produce other reactive species such as hydrogen peroxide, but *R. equi* is highly resistant to hydrogen peroxide stress (Meijer and Prescott, 2004). In fact, hydrogen peroxide triggers the induction of *vap* genes of the organism especially *vapA* and *vapG* (Benoit *et al.*, 2002; Meijer and Prescott, 2004).

Though the exact mechanism of protective immune response in adult horses is unknown, there is evidence that pulmonary clearance of *R. equi* is as a result of a Th1 response where virulent strains are cleared with an increase in $CD4^+$ and $CD8^+$ T-lymphocytes (Hines *et al.*, 2001). Experiments have also shown that T-lymphocytes, obtained from lungs of adult horses following immune clearance, proliferate in culture in response to *R. equi* antigens with the production of IFN-γ but not interleukin-4 (Lopez *et al.*, 2002).

Antibody-mediated responses also play a significant role in protective immunity. In foals, colostrum-derived maternal antibody declines to its lowest level by 8 weeks of age, after which time they start to produce antibodies to the organism. Foals with low levels of maternal antibody are particularly susceptible to *R. equi* infections (Prescott, 1991). Immunoglobulin G (IgG) plays an important role in opsonisation and phagocytosis of virulent *R. equi* while mucosal IgA (from infected foals) is highly reactive to a specific B-cell epitope on VapA (Muscatello, 2012). The apparent absence of IgA from the nasal mucosa (during the first 4 weeks of life) and

the low circulating antibody titres (usually between 1 and 3 months) may increase the susceptibility of foals to *R. equi* infection (Muscatello 2012). Adult horses are generally resistant to *R. equi* infection since most of them have measurable antibody titres to the organism thus are said to be hyperimmune. This indicates the inverse relationship between antibody levels and disease prevalence and severity (Hietala *et al.*, 1985). Results with passive immunisation of hyperimmune sera to foals exposed to *R. equi* protected them from disease (Mosser and Hondalus, 1996). In addition to humoral immunity, cell-mediated immunity could also be transferred by mare's colostrum to foals (Porto *et al.*, 2014).

1.8 Mechanism of lipoprotein processing

Lipoproteins in monoderm Gram-positive bacteria are cell envelope proteins that are anchored onto the outer leaflet of the plasma membrane (Hutchings *et al.*, 2009). They contain lipids covalently attached to an *N*-terminal cysteine residue and this acylated *N*-terminal cysteine motif is believed to anchor the proteins to the plasma membrane *via* hydrophobic interactions of the acyl groups (Kovacs-Simon *et al.*, 2011). Hence, the insertion of the lipid group provides a mechanism for localising proteins to bacterial cell membranes (Sutcliffe *et al.*, 2010). In Gram positive bacteria (including *R. equi*), lipoproteins perform diverse functions such as antibody resistance, substrate binding, signalling, conjugation and secretory functions (Sutcliffe and Russel, 1995). These proteins (especially of intracellular pathogens like *R. equi*) are also primary targets of the host immune response thus they may be key molecules to induce immune protection, and consequently, key targets for vaccine development (Barbey *et al.*, 2009). The preprolipoproteins (especially those of the membrane-anchored lipoproteins) contain an *N*-terminal signal peptide (or

signal sequence) that is involved in the targeting of proteins towards organelles or in their secretion which require the translocation of such proteins across membranes (Creuzenet *et al.*, 1997). The signal sequence is the part that is recognised by the receptors of the translocation machinery, serves as a topological determinant for the prolipoprotein in the membrane and as a signal to inhibit the folding of nascent chains. This avoids the activation of potentially harmful secretory enzymes inside the cell and concurrently retaining translocation competence (Van Roosmalen *et al.*, 2004). A signal peptide is usually 14 – 25 amino acids long and consists of three distinct domains: the amino (N-), hydrophobic (H-), and the carboxylic-terminal (C-) regions (Figure 1.15). The H-region is the longest region of the signal peptide and consists of 10 – 15 amino acids. This region tends to organise into an α-helical conformation when in contact with a membrane lipid phase because of the presence of hydrophobic residues (Van Wely *et al.*, 2001). At the *N*-terminal side of the H-region, there is a stretch of positively-charged amino acids which is the N-region of the signal peptide. This region determines the orientation of the signal peptide due to its relatively high positively-charged polar nature (Van Roosmalen *et al.*, 2004). The C-region contains the cleavage site (lipobox) for the signal peptidase (Lsp) and is hydrophilic in nature (Van Wiley *et al.*, 2001).

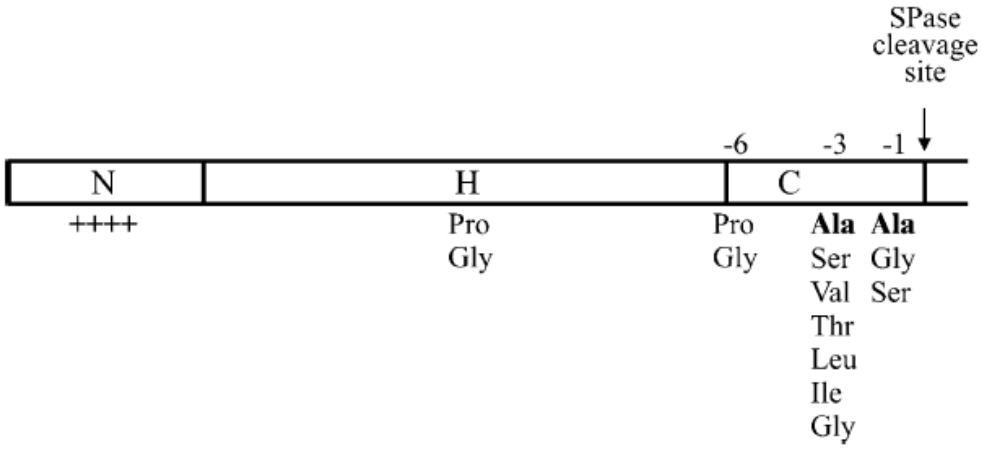


Figure 1.15 Structure of a typical bacterial signal peptide.
It consists of a positively charged N-terminus (N-region), a central hydrophobic region (H-region), and a polar C-terminal region (C-region). Helix-breaking proline or glycine residues are often found in the middle of the H-region and between the H- and C-regions at the −6 position relative to the cleavage site. The Lsp recognition sequence consists of small aliphatic residues at positions −1 and −3, relative to the cleavage site. The most common residue at these positions is Ala (Adapted from Van Roosmalen *et al*., 2004).

This conserved motif directs the protein to the lipoprotein biogenesis machinery after transport (Hutchings *et al.*, 2009). The signal peptide is then cleaved in a reaction catalysed by lipoprotein signal peptidase II (LspII) (for Gram positive bacteria) leaving the lipid-modified cysteine as the *N*-terminus of the mature lipoprotein (Sutcliffe and Harrington, 2002) (Figure 1.16). In many bacteria, the ligation of the diacylglycerol to the cysteine moiety within the lipobox by the prolipoprotein diacylglyceroltransferase (lgt) is a pre-requisite for the activity of Lsp (Sander *et al.*, 2004; Denham *et al.*, 2009). Structural studies reveal that in some bacteria, the diacylglycerol is linked to the *N*-terminal cysteine *via* a thioether bond and this lipid group serves to orientate the protein by anchoring it to the inner leaflet of the membrane (Sutcliffe and Harrington, 2002). The diacylglycerol group of the

lipoproteins are derived from the membrane phospholipids which may not only provide anchorage of the lipoprotein to the membrane but also other functional roles since the lipoproteins are surface exposed (Kovacs-Simon *et al.*, 2011).

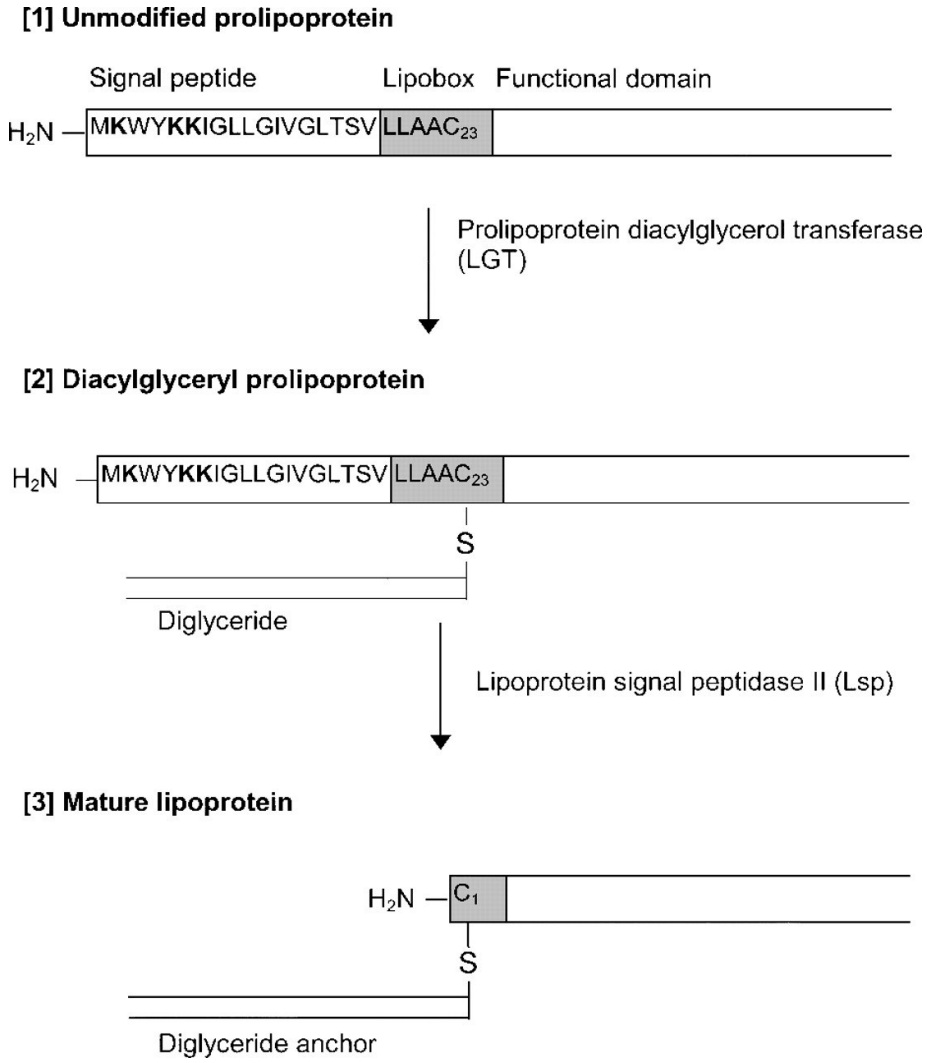


Figure 1.16 Lipoprotein biosynthesis in Gram-positive bacteria.
Lipid modification of prolipoproteins (by thioether linkage of diglyceride to the lipobox cysteine) is followed by signal peptide cleavage, generating the mature lipid-anchored lipoprotein. The signal sequence shown is just a representation of a classical signal peptide (Adapted from Sutcliffe and Harrington, 2002).

Apart from VapF, VapI, and VapX, the full length proteins of the Vap family exhibit significant homology in their C-terminal regions (Figure 1.17). Even though there are significant differences in the amino termini of the proteins of the Vap family (VapA, -B, -C, -D,-E, -G, -H, -J, -K), they all possess a predicted signal peptide with an *N*-terminal, hydrophobic, and a C-terminal regions with a potential cleavage sites for signal peptidase (Byrne *et al.*, 2001; Takai *et al.*, 2000). This clearly shows that they possess features of secretory proteins. VapA has long been recognised as a lipidated protein but the lack of the characteristic cysteine residue employed in lipoprotein biogenesis suggests that this protein might possess an unusual Lgt-independent lipid modification.

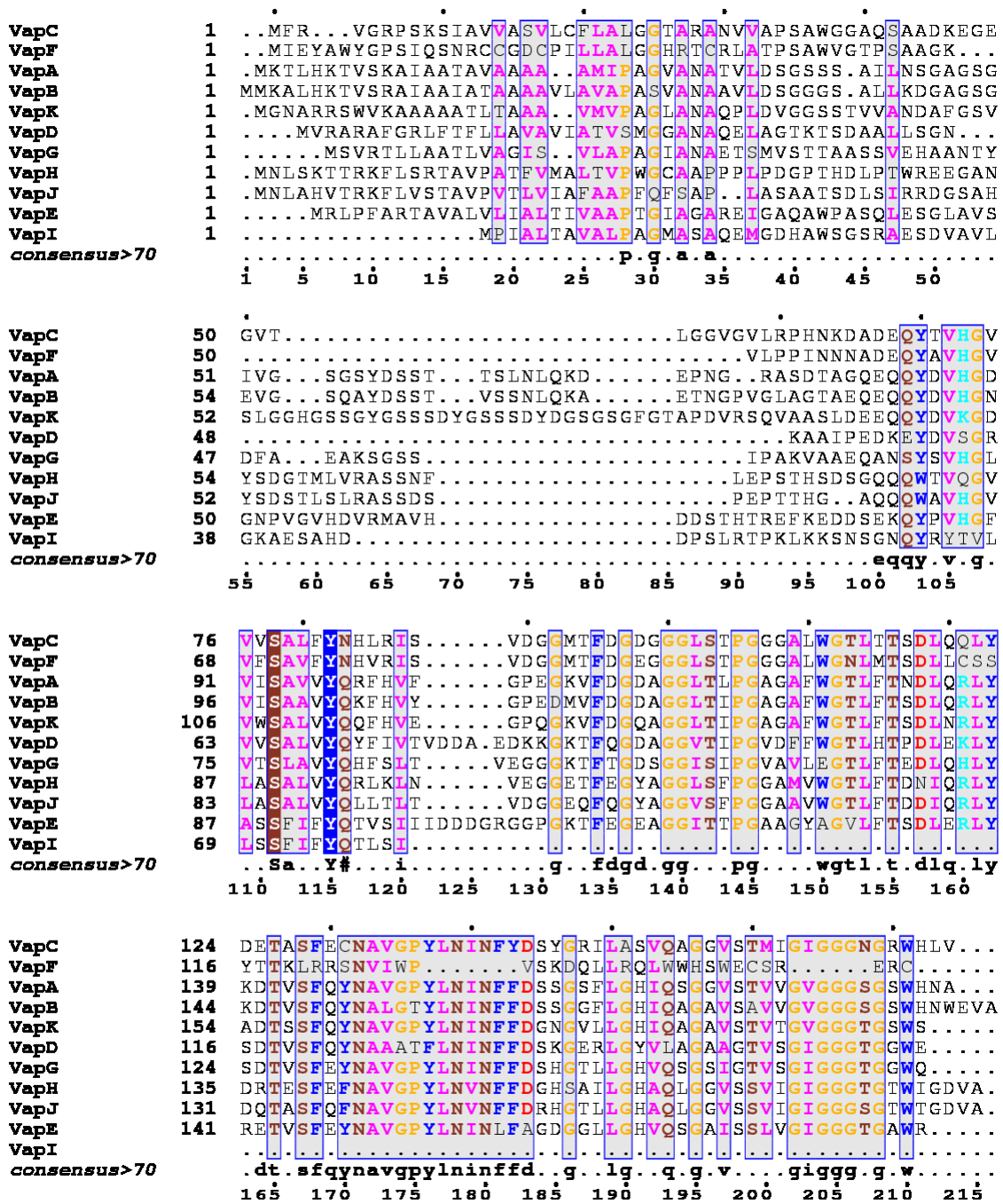


Figure 1.17 Sequence alignment of the Vap proteins from the vapB plasmid pVAPB1593 and the vapA plasmid pVAPA1037

Alignment was done using ClustalW 2.1 multiple alignment programme formatted using ESPript3.0 (Gouet *et al.*, 2003). Apart from VapF and VapI, the proteins exhibit significant homology in their C-terminal regions.

1.9 Aim of the study

The virulence of *R. equi* is largely attributed to its virulence plasmid, which bears a pathogenicity island that encodes VapA (a 15 – 17 kDa immunodominant surface-expressed protein). Even though the exact mechanism of virulence is not known, VapA is thought to mediate the intracellular growth of the organism in macrophages and disease development *via* the arrest of phagosome maturation (Polidori and Haas, 2006; Barbey *et al.*, 2009; Letek *et al.*, 2010).

VapA is susceptible to trypsin digestion and can be labelled with a biotin-containing bifunctional agent which suggests that it is a surface-anchored protein. Biotin esters do not penetrate the cell membrane thus have been used as biotinylation reagents for membrane proteins (Smither *et al.*, 2007). Its metabolic labelling with [^{3}H] palmitate shows that it is a lipid-modified protein, a key property for cell wall anchorage (Meijer and Prescott, 2004). In addition, VapA has a predicted signal sequence, a positively charged amino acid region and a hydrophobic region with potential cleavage site for signal peptidase at the C-region (Bryne *et al.*, 2001). These properties show that VapA is a lipoprotein that maybe non-covalently attached to the rhodococcal cell envelope. Even though VapA is an exported secretory protein, nucleotide sequence analysis studies also show that preproVapA has no lipobox and the mature VapA lacks the typical acylated *N*-terminal cysteine motif (Tan *et al.*, 1995) which implies that *Rhodococcus equi* may possess a different mechanism to lipidate its proteins especially VapA *i.e.* independent of prolipoprotein diacylglycerol transferase (Lgt) (Meijer and Prescott, 2004). Reports have it that VapA is an acylated protein and it has been suggested that lipid modification of VapA may likely occur at a site close to the *N*-terminus of the mature protein (Tan *et al.*, 1995). Thus

one of the focuses of the study will be the investigation of the post-translational lipidation status of VapA.

In *Corynebacterium* (a related genus of *Rhodococcus*), channel forming porins (e.g CgPorA in *Corynebacterium glutamicum*) are modified by *O*-linked mycolic acids and this post-translational mycoloylation enhances their pore-forming abilities (Huc *et al.*, 2010). Since lipoproteins and channel forming porins are non-covalently attached to the wall, one objective of this research is to establish whether VapA has mycolyl residues analogous to CgPorA in *Corynebacterium glutamicum* since in some cases the mycolic acid layer serves as a potential anchoring point for lipoproteins (Kovacs-Simon *et al.*, 2011). The aim was to investigate this hypothesis *via* chemical analysis of purified material and heterologous expression of cloned *vapA* in *C. glutamicum* to establish whether the unusual modification of the protein can be generated in an organism of a related genus. *Corynebacterium glutamicum* could represent a suitable model for functional studies of expression of rhodococcal genes related to cell wall function because it possesses the simplest cell wall structure of the mycolata and hence is functionally and structurally close to *R. equi*. It represents a good model organism as it is a fast grower and does not form clumps and heterologous expression of antigens of related bacteria had already proved to be effective (Puech *et al.*, 2001).

Expression of *vapA* in either an isogenic plasmidless derivative or a *vap* PAI deletion mutant showed that VapA alone, in the absence of other vap PAI products, is not sufficient to restore the capacity of *R. equi* to proliferate in macrophages or mouse organs and to colonise the lungs of experimentally infected foals (Vazquez-Boland *et al.*, 2010; Vazquez-Boland *et al.*, 2013). Hence the Vap proteins could interact to

sustain complete virulence of the organism in the host. Potential interaction of Vaps was to be investigated using biophysical techniques including surface plasmon resonance and analytical ultracentrifugation.

In the absence of mechanistic insight, a structural study of the virulence associated proteins using X-ray crystallography was also planned. Vap genes were to be cloned and expressed in *E. coli*. The proteins would be purified to commence the production of crystals for X-ray diffraction studies, which might deliver structural information that might inform subsequent functional studies.

CHAPTER TWO

2 Materials and methods

2.1 Chemicals and Reagents

All chemicals and reagents were of analytical grade hence were used without further purification. Solutions and buffers were prepared using Milli-Q (18 $m\Omega.cm^{-1}$) water and stored at room temperature unless otherwise stated. All primers for PCR were supplied by Eurofins Genomics, those used for sequencing were synthesised and used at GATC Biotech.

2.2 Media

Media were prepared using distilled water, sterilised by autoclaving and stored at room temperature except for solid media that were stored at 4°C.

2.2.1 Liquid media (Broth)

Liquid media (Broth) used in the study include Luria-Bertani (LB), Terrific Broth, and Brain Heart Infusion (BHI). Recipes for each are given in appendix A.

2.2.2 Solid media

Solid media used were LB agar and BHI agar. Recipes are given in appendix A.

2.2.3 Selective media (Media containing antibiotic)

Antibiotics used for the preparation of selective media are ampicillin and kanamycin. Preparations of selective media are given in appendix A.

2.3 Bacterial strains and plasmids

Bacterial strains used include *Rhodococcus equi* 103^{+} (I. C. Sutcliffe collection), *Rhodococcus equi* Δ*kasA* (Sydor *et al*., 2013), *Corynebacterium glutamicum* ATCC 13032 (ATCC), *E. coli* Top 10 (Invitrogen) and *E. coli* BL21 (DE3) (Invitrogen). *E. coli* Top 10 was used as host for cloning while *E. coli* BL21 (DE3) was used for recombinant gene expression. Bacterial strains were stored in 20 % glycerol (sterile) and maintained at -80°C while plasmids were stored at -20°C. The plasmids used include pET23a (Novagen), pET28a (Novagen), *E. coli-C. glutamicum* shuttle vector pEKEx2 (Eikmanns *et al*., 1991) and the cosmid p10.1 (Raul Miranda-Casoluengo, University College Dublin). The relevant features of pET23a, pET28a, and pEKEx2 are briefly described in appendix B.

2.4 Production of chemically competent cells

Chemically competent cells were prepared according to the method of Sambrook and Russell (2001) with some modifications. Briefly, 5 mL of LB broth was inoculated with a single colony of *E. coli* Top10 cells and grown overnight at 37°C with shaking at 200 rpm as a starter culture. One millilitre of the starter culture was transferred to 50 mL of LB broth and grown at 37°C with shaking at 200 rpm until an OD_{600} of 0.35 – 0.4 was obtained. Cells were recovered by centrifuging 25 mL of the culture at 2880 x *g* for 10 min at 4°C. The supernatant was discarded and the pellet re-suspended in 15 mL of ice-cold universal solvent (0.08 M $MgCl_2$ solution containing 0.02 M $CaCl_2$ and autoclaved), held on ice for 60 min and recovered by centrifugation as before. The supernatant was discarded and the pellet re-suspended in 1 mL of cold $CaCl_2$ (0.1 M). The cells (50 µL portion) were used directly for transformation or re-suspended in glycerol (15 % - 20 % final concentration) and

dispensed as 50 µL aliquots in sterile 1.5 mL microcentrifuge tubes which were flash frozen in liquid N_2 and stored at -80°C.

2.5 Plasmid maps

The maps of the plasmids (where applicable) were constructed using PlasMapper which is a web-based plasmid map drawing tool developed by Dong *et al*. (2004). The plasmid sequence was fed into programme which generated the map.

2.6 Estimation of Protein and DNA concentrations

The concentration of DNA and purified proteins were quantified using a NanoDrop 1000 Spectrophotometer following the manufacturer's protocol. The estimation for other protein solutions (where applicable) was made *via* the Bradford reagent using bovine serum albumin (BSA) as the standard. A standard curve (Appendix C) was generated with different concentrations of BSA (0.1 mg/mL – 1.0 mg/mL) and the concentration of the protein solution (treated identically to the standards) was determined with reference to the standard curve. Briefly, 0.1 mL of either standard or sample was delivered in a cuvette. Bradford reagent (3 mL) was added to the protein solution, mixed using a pipette and held at room temperature for 10 min before measuring absorbance at 595 nm.

2.7 DNA methods

2.7.1 Extraction of Corynebacterium glutamicum genomic DNA

A single colony of *Corynebacterium glutamicum* ATCC 13032 was picked from a culture plate and grown in 25 mL of BHI broth at 30°C to OD_{600} of 0.6 to 0.8. The cells were recovered by centrifuging culture at 2880 x *g* for 30 mins at 20°C and re-

suspended in 450 µL of GTE-RNase buffer (Appendix D) and transferred into a 1.5 mL microcentrifuge tube. Fifty microlitres of lysozyme (10 mg/mL) was added to the mixture and incubated overnight at 30°C. The next day, 100 µL of sodium dodecyl sulphate (10 % w/v), 20 µL of proteinase K (15 mg/mL) were added, mixed gently and held at 55°C for 3 hrs. The mixture was allowed to cool and 10 µL of RNase A (10 mg/mL) was added, mixed and incubated for 30 min at 37°C. Two hundred microlitres of NaCl (5 M), 1 mL of chloroform: isoamyl alcohol (24:1) were added, mixed and centrifuged at 13,000 rpm (using a bench top centrifuge) for 10 min. The upper aqueous layer was collected and delivered into a new microcentrifuge tube and mixed with 1 mL of chloroform: isoamyl alcohol (24:1) and centrifuged. The aqueous layer was recovered and mixed with 0.7 volume of ice-cold isopropanol (about 0.7 mL) by inversion, and centrifuged at 4°C, 27,000 x *g* for 30 min. The supernatant was removed and the pellet was washed with 0.7 mL of ice-cold ethanol (70 %) and centrifuged at 4°C, 27,000 x *g* for 30 min. The ethanol was removed with a pipette and the pellet was allowed to dry at room temperature on the bench. The pellet was finally re-suspended in 20 µL of elution buffer from QIAprep spin mini-prep kit (Qiagen).

2.7.2 Extraction of plasmid DNA

Plasmid DNA was extracted using a QIAprep spin mini-prep kit (Qiagen) following the manufacturers' protocol. Recombinant plasmids were sent for sequence analysis thereafter.

2.7.3 Electrophoretic analysis of DNA

The gel casting tray was prepared by boiling 0.3 g of dry agarose powder in 30 mL of Gel buffer (Tris-Acetic acid-EDTA or TAE buffer, appendix D). A comb was inserted into the electrophoresis tray and the molten agarose solution poured in a casting chamber and allowed to set. The casting chamber was placed in the electrophoresis chamber and covered with TAE buffer up to the maximum mark in the tank. The comb was removed and sample (1 µL) was loaded on each well after mixing with 1 µL of DNA loading buffer (Appendix D). The gel was run at a constant current of 120 mA for 25 min. The gel was placed in a stain box containing ethidium bromide solution (0.5 µg/mL) for 10 min and visualised under UV light and an image captured using a Bio-Rad Gel Doc system with Quantity One software.

2.7.4 Oligonucleotide primers for gene cloning

The sequences that code for the virulence associated proteins A, C, D, E, G, H and corynomycolyl transferase (Cmt) were amplified via PCR. Since the proteins contain predicted signal peptides, oligonucleotide primers were designed for the amplification of the genes in such a way to avoid the regions which code for the putative signal peptides (for cloning expression in *E. coli*) as validated by SignalP (Appendix E). Primers (for ligation and restriction digestion) were designed to have restriction enzyme sites at the 5′ and 3′ ends and addition number of bases at each end to allow for efficient binding and cleavage. The genes were amplified using oligonucleotide primers (as shown in table 2.1) for insertion of the ORFs into either pET23a or pET28a (in frame for expression in *E. coli*) or pEKEx2 (in frame for expression in *C. glutamicum*).

Table 2.1 Oligonucleotides used for the study.

OLIGO NAME	SEQUENCE (5′- 3′)	Restriction endonuclease
pET23/28-VapAup	GATCGATC**CATATG**ACCGTTCTTGATTCCGGTAGCAGCAGTGCGGTGCG	NdeI
pET23-VapAfull	GATCGATC**CATATG**AAGACTCTTCACAAGACGGTTTCTAAGGCG	NdeI
pET23-VapAlo	TATAAATA**GCGGCCGC**GGCGTTGTGCCAGCTACCAGAGCCG	NotI
pET28-VapAlo	TATAAATA**GCGGCCGC**CTAGGCGTTGTGCCAGCTACCAGAGCCG	NotI
pET23/28-VapCup	GATCGATC**CATATG**AATGTAGTCGCTCCGTCGGCGTGG	NdeI
pET23-VapClo	AATTTATA**GCGGCCGC**CACCAAATGCCATCGCCCATTCC	NotI
pET28-VapClo	AATTTATA**GCGGCCGC**TCACACCAAATGCCATCGCCCATTCC	NotI
pET23/28VapDup	GTAGGAAC**CATATG**CAGGAGCTAGCTGGCACCAAGAC	NdeI
pET23-VapDlo	GTAGGAAC**GAATTC**CTCCCACCCGCCAGTG	EcoRI
pET28-VapDlo	GTAGGAAC**GAATTC**CTACTCCCACCCGCCAGTG	EcoRI
pET23/28-VapEup	GTACGATC**CATATG**CGGGAAATTGGCGCCCAG	NdeI
pET28-VapElo	GTACGATC**GAATTC**CTATCGCCAGGCGCCAGTG	EcoRI
pET23/28-VapGup	GACGATC**CATATG**GAAACTTCAATGGTATCCACTACAGCAGCATCG	NdeI
pET-28VapGlo	GACGATC**GAATTC**CTATTGCCACCCTCCGGTTCC	EcoRI
pET-23/28VapHup	GACGATC**CATATG**GCCCCGCCACCATTACCAGATG	NdeI
pET28-VapHlo	GACGATC**GAATTC**CTACGCTACATCGCCTATCCAGGTTCCCGTAC	EcoRI
pET23/28-Cmtup	AATCAATT**CATATG**GCCGAAGTAACCCCAGCAGACG	NdeI
pET23-Cmtlo	AATCAATC**GCGGCCGC**GGCCTCTAGCTCAAACGC	NotI
pET28-Cmtlo	AATCAATC**GCGGCCGC**CTAGGCCTCTAGCTCAAACG	NotI
pEKEx2VapAup	GATCGATC**GGTACCAAGGAGATATAGAT**ATGAAGACTCTTCACAAGACGG	KpnI
pEKEx2-VapAlo	GATCGATC**GAATTC**CTAGGCGTTGTGCCAGCTAC	EcoRI
pEKEx2-VapAHislo	GATCGATC**GAATTC**CAACTCAGCTTCCTTTCGGGCTTTG	EcoRI

All oligonucleotides were designed to introduce recognition sites for restriction endonucleases upstream and downstream of the target sequence (recognition sites are underlined, and restriction endonucleases are in brackets) to facilitate cloning into vectors. Highlighted in bold (but not underlined) is the ribosomal binding site.

PCR reactions used Vent DNA polymerase (New England Biolabs) which has a high proof reading ability. The cosmid p10.1 was used as a template for the amplification of *vapA*, *vapC*, and *vapD* (for cloning into either pET23a or pET28a) as it contains coding sequences for the respective Vap proteins. For the amplification of *vapE*, *vapG*, and *vapH*, relcalcitrant genes (which codes for the VapE, VapG and VapH) were sourced from Eurofins MWG Operon while for *cmt*, *C. glutamicum* genomic DNA was used as the template. For each oligonucleotide pair, reaction mixture (50

µL final volume) was made up of Thermopol buffer (1 time strength), deoxynucleotide triphosphates mix (80 µM each), oligonucleotide primer (20 pg/µL each), 1 – 5 µg of template DNA and 20 units of Vent DNA polymerase. The thermal cycling conditions are shown in table 2.2.

Table 2.2 Thermocycler program for PCR

Phase	Initial denaturation	Amplification			Polishing
		Denaturation	Annealing	Extension	
T°C	95	95	55	72	72
Time (mins)	3	0.5	0.5	1 / kb plasmid	10
Cycles (n)	1		35		1

* For ORFs less than 500 bp, extension time was set at 30 seconds

2.7.5 Gel purification of amplified products

Where appropriate, samples containing PCR products were pooled together and loaded on a broad well in an agarose gel. After running the gel, it was stained in freshly prepared ethidium bromide solution (0.5 µg/mL) and visualised using the prep UV setting of a trans-illuminator. The band corresponding to DNA size of interest was excised from the gel and DNA purified using either QIAquick gel extraction kit or NzyTech Gelpure kit following the manufacturers' protocols. A sample was analysed for purity using agarose gel electrophoresis.

2.7.6 Cloning of amplified DNA in pET23a and pET28a

The genes were cloned into either the NotI/NdeI or Nde1/EcoRI cloning sites of both pET23a and pET28a vectors. Purified PCR products (DNA inserts) for *vapA*, *vapC* and *cmt* were separately digested with NdeI and NotI restriction enzymes to generate cohesive ends in NE Buffer 3 which supports maximal activity for digestion using

those two enzymes. While purified PCR products of *vapD*, *vapE*, *vapG*, and *vapH* were digested separately with NdeI and EcoR1 in NE buffer 4 as it supports maximum activity for the enzyme pair. The vectors were also digested with either NotI/NdeI or NdeI/EcoRI. The mixture for restriction digestion is shown in table 2.3.

Table 2.3 Restriction enzymes digest mixture

Item	Volume (µL)
Plasmid or PCR product	40
NE Buffer (10 x strength)	10
BSA (100 x strength)	1
Restriction enzyme 1	5
Restriction enzyme 2	5
Sterile Milli Q water	39

Each digestion mixture was incubated at 37°C for 3 hrs. Plasmids (now linearised) and DNA insert were gel purified thereafter. Each insert was ligated to linearised plasmid in a reaction whose constituents were 1 µL of gel-purified linearised plasmid DNA solution, 11.5 µL of insert DNA solution (*vap or cmt*), 1.5 µL of T4 DNA ligase buffer (10 x strength) and 1 µL of T4 DNA ligase (400,000 U/mL). Each mixture was incubated at 16°C overnight and used to transform chemically competent *E. coli* Top10.

2.7.7 Cloning of amplified DNA into pEKEx2

In this case, a second gene amplification was performed using pET23-vapAfull as the template with the primer pairs pEKEx2-vapAup/pEKEx2-vapAlo and pEKEx2-vapAup/pEKEx2-vapAhislo as shown in table 2.1. After amplification of the targets, the products were gel purified, restriction digested with *Kpn*I and *EcoR*I and ligated to pEKEx2 (which was also restriction digested with *Kpn*I and *EcoR*I, and gel purified).

2.7.8 Transformation of chemically competent E. coli

Fifteen microlitres of suspected ligated DNA was added to a 50 µL aliquot of freshly thawed glycerol stock of chemically competent *E. coli* Top10 and incubated on ice for 20 min. It was heat shocked at 42°C in a water bath for 2 min and transferred immediately to ice for another 2 min for recovery. LB broth (200 µL) was added to the solution and incubated for 45 min at 37°C. A 200 µL sample of the suspension (containing transformed cells) was spread on LB agar (supplemented with 100 µg/mL ampicillin for cloning into pET23a or 25 µg/mL kanamycin for cloning into pET28a/pEKEx2) and incubated at 37°C overnight. Single colonies (transformants) were selected and inoculated into 5 mL of LB broth (with either ampicillin or kanamycin) and incubated overnight at 37°C with shaking at 150 rpm. The broth was used for plasmid extraction.

2.7.9 Electroporation of pEKEx2-derived plasmids into Corynebacterium glutamicum

Electroporation of pEKEx2-*vapA* and pEKEx2-*vapAhistag* into *C. glutamicum* was performed according to the method of Van der Rest *et al*. (1999) with modifications. Ten millilitres of LB broth (containing 2 % glucose sterile filtered) was inoculated with a colony of *C. glutamicum* from a fresh plate as a starter culture overnight at 30°C with shaking at 150 rpm. One hundred millilitres of special medium (for growing electro-competent cells, appendix A) was inoculated with 5 mL of the overnight culture and grown to an OD_{600} of about 1 at 18°C and 150 rpm. The culture was chilled on ice and pelleted by centrifuging at 2880 x *g* for 10 min at 4°C. The pellet was re-suspended and washed three times with 20 mL ice-cold Tris-Glycerol buffer (1 mM Tris-HCl buffer, pH 7.5; 10 % v/v glycerol) by centrifuging

at 2880 x *g* for 10 min at 4°C. The pellet was re-suspended in 1 mL ice-cold 10 % glycerol (v/v) and stored as aliquots in 200 µL volumes in pre-chilled 1.5 mL microcentrifuge tubes and stored at -80ºC (cells now electro-competent). Electro-competent cells (150 µL) and 2 µL of plasmid DNA (recombinant) were transferred to a pre-chilled electroporation cuvette (0.2 cm gap), topped gently with 0.8 mL of ice-cold 10 % glycerol and electroporated (using a Gene Pulser System from Bio-Rad) with a single pulse with parameters set at 25 µF, 2500 V, and 600 Ω. Solution was transferred immediately into 4 ml of pre-warmed (at 46ºC) BHI medium containing 2 % sorbitol (BHIS) and heat shocked by incubating for 6 min at 46°C in a water bath, and immediately for 1 hr at 30°C for recovery. Two hundred microlitres of the mixture was subsequently plated on BHIS agar (containing 25 µg/mL kanamycin) and incubated at 30°C overnight.

2.7.10 Site-directed mutagenesis

The plasmid pEKEx2-*vapA* was amplified using primers that were designed to bear point mutations in *vapA*. The oligonucleotides used for site-directed mutagenesis are shown in Table 2.5.

PCR reactions used pEKEx2-*vapA* as the template, primers (table 2.4), Vent DNA polymerase, thermopol reaction buffer, dNTPs, and distilled water with volumes and concentrations as described in section 2.7.4. However the thermocycler programme is shown in table 2.4.

Table 2.4 Thermocycler program for site-directed mutagenesis

Phase	Initial denaturation	Amplification			Polishing
		Denaturation	Annealing	Extension	
T°C	95	95	55	68	68
Time (mins)	3	0.5	1	1 minute/kb plasmid length	30-60 minutes
Cycles (n)	1		12		1

After the reaction, the tube was later placed on ice and 1 µL of the restriction enzyme *Dpn*I (20,000 U/mL) was added to the mixture, gently mixed by pipetting and incubated for 1 hr at 37°C to digest parental DNA. One microlitre of the solution was later used to transform chemically competent *E. coli* Top 10 cells. A single transformant was grown in 5 mL LB broth (containing kanamycin), plasmid extracted and sent for sequence analysis. The pEKEx2-*vapA*mutant was later used to transform electro-competent *C. glutamicum*.

Table 2.5 Oligonucleotides used for site-directed mutagenesis.
The mutant sites are highlighted in bold. Where the targets are not continues (in more than once codon), the codons in the antisense sequence (per primer pair) are underlined

OLIGO NAME	SEQUENCE (5′- 3′)
VapAST/ABlock1up	GGCTGCGCTAATGCG**GCC**GTTCTTGATTCCG
VapAST/ABlock1lo	CGGAATCAAGAACGG**CCG**CATTAGCGCAGCC
VapAST/ABlock2up	GCTAATGCGACCGTTCTTGAT**GCC**GGTAGCAGCAGTGCGATTCTC
VapAST/ABlock2lo	GAGAATCGCACTGCTGCTACC**GGC**ATCAAGAACGGTCGCATTAGC
VapAST/ABlock3up	CCGTTCTTGATTCCGGT**GCCGCCGCT**GCGATTCTCAATAGTGGG
VapAST/ABlock3lo	CCCACTATTGAGAATCGC**AGCGGCGGC**ACCGGAATCAAGAACGG
VapAST/ABlock4up	GCAGTGCGATTCTCAATGCTGGGGCAGGC**GCT**GGCATTGTCGGTTCTGGG
VapAST/ABlock4lo	CCCAGAACCGACAATGCC**AGC**GCCTGCCCCAGCATTGAGAATCGCACTGC
VapAST/ABlock5up	GGCATTGTCGGTGCTGGG**GCC**TATGACAGCTCG
VapAST/ABlock5lo	CGAGCTGTCATA**GGC**CCCAGCACCGACAATGCC
VapAST/ABlock6up	CTGGGAGCTATGAC**GCCGCGGCGGCTGCG**TTAAACCTTCAG
VapAST/ABlock6lo	CTGAAGGTTTAACG**CAGCCGCCGCGGCGT**CATAGCTCCCAG
VapAST/ABlock7up	GAACGGTCGAGCAGCCGAT**GCC**GCCGGGCAAGAG
VapAST/ABlock7lo	CTCTTGCCCGGC**GGC**ATCGGCTGCTCGACCGTTC
VapAST/ABlock8up	CACGGAGACGTCATC**GCC**GCGGTCGTCTACCAG
VapAST/ABlock8lo	CTGGTAGACGACCGC**GGC**GATGACGTCTCCGTG
VapAST/ABlock9up	GATGCAGGGGGACTC**GCG**CTTCCTGGGGCCGGC
VapAST/ABlock9lo	GCCGGCCCCAGGAAG**CGC**GAGTCCCCCTGCATC
VapAST/ABlock10up	GTTCTGGGGGGCTCTCTTC**GCA**AATGACCTTCAG
VapAST/ABlock10lo	CTGAAGGTCATT**TGC**GAAGAGAGCCCCCCAGAAC
VapAST/ABlock11up	CTCTACAAAGACGCCGTC**GCG**TTCCAGTACAAC
VapAST/ABlock11lo	GTTGTACTGGAA**CGC**GACGGCGTCTTTGTAGAG
VapAST/ABlock12up	CAACTTCTTCGATGCCGCAGGT**GCC**TTCCTCGGCCATATC
VapAST/ABlock12lo	GATATGGCCGAGGAA**GGC**ACCTGCGGCATCGAAGAAGTTG
VapAST/ABlock13up	CTCGGCCATATCCAG**GCC**GGTGGAGTTAGTACTGTG
VapAST/ABlock13lo	CACAGTACTAACTCCACC**GGC**CTGGATATGGCCGAG
VapAST/ABlock14up	CAGTCCGGTGGAGTT**GCTGCT**GTGGTGGGCGTCGGCGTC
VapAST/ABlock14lo	GACGCCGACGCCCACCAC**AGCAGC**AACTCCACCGGACTG
VapAST/ABlock15up	GGCGTCGGCGGCGGCGCTGGT**GCC**TGGCACAACGCCTAG
VapAST/ABlock15lo	CTAGGCGTTGTGCCA**GGC**ACCAGCGCCGCCGCCGACGCC

2.8 Protein Methods

2.8.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

A discontinuous electrophoresis gel system (Laemmli, 1970) was used with a 12 % resolving gel in a mini-gel format (Miniprotean® II, Bio-Rad). The formulation of the buffers required (resolving gel buffer, stacking gel buffer, sample loading buffer, and SDS running buffer) and staining solutions are detailed in appendix D. The resolving gel was mixed as detailed in Table 2.6 and was poured immediately after preparation into the gel casting cassette, overlaid with water and allowed to set for about 20 min. The water was removed and the stacking gel (Table 2.7) was poured immediately after preparation into the casting cassette on top of the resolving gel, the comb was inserted and allowed to set. The gel sandwich was clamped and placed inside the chamber, which was filled with SDS running buffer. Each protein sample (including standards) was mixed with loading buffer in a volumetric ratio of 5:1, and held at 95°C for 3 min and allowed to cool. The comb was removed and 10 µL of protein sample was loaded. Gels were electrophoresed for 50 min at a constant voltage of 200 V. The gel was removed, immersed in Coomassie blue stain for 15 min, rinsed with water and soaked in de-stain overnight. Gels images were captured (after washing the gel in water) using a Gel Doc system with Bio-Rad Quantity One software.

Table 2.6 SDS-PAGE resolving gel recipes

Ingredient	Volume (mL)
Resolving gel buffer (1.5 M Tris.HCl, pH 8.8, 4% SDS)	2.50 mL
40% Acrylamide:Bisacrylamide, (37.5: 1)	3.00 mL
Distilled water	4.50 mL
*Ammonium per sulphate (10 %)	0.05 µL
TEMED	0.01 µL

TEMED is *N,N,N′,N′*-Tetramethylethylenediamine. *Ammonium persulphate was freshly prepared before use and added last.

Table 2.7 SDS-PAGE Stacking gel recipes

Ingredient	Volume (mL)
Stacking gel buffer (0.5 M Tris.HCl, pH 6.8, 4 % SDS)	1.00
40% Acrylamide:Bisacrylamide (37.5: 1)	0.50
Distilled water	2.50
*Ammonium per sulphate (10 %)	0.03
TEMED	0.01 µL

TEMED is *N,N,N′,N′*-Tetramethylethylenediamine. *Ammonium persulphate was freshly prepared before use and added last

2.8.2 Non-denaturing (Native) polyacrylamide gel electrophoresis (Native-PAGE)

This was also used with a 12 % resolving gel in a mini-gel format (Miniprotean® II, Biorad) as reported in SDS-PAGE but the stacking gel buffer, resolving gel buffer, running buffer and the sample loading buffer did not contain sodium dodecylsulphate. The protein and sample buffer were mixed and immediately loaded onto the gel without boiling. Staining and de-staining were done as reported for SDS-PAGE.

2.8.3 Urea polyacrylamide gel electrophoresis (Urea-PAGE)

A 12 % resolving gel in a mini-gel format (Miniprotean® II, Bio-Rad) was also used but the resolving and stacking gels were supplemented with urea to a final concentration of 1 %. Thus the recipes for Urea-PAGE are given in tables 2.8 and 2.9. Samples were mixed with loading buffer (0.2 M Tris.HCl, pH 6.8, containing 75 % glycerol, 3% bromophenol blue) and loading unto the wells without boiling. Running buffer was the same as the Native-PAGE.

Table 2.8 Urea-PAGE resolving rel casting recipes

Ingredient	Volume (mL)
Resolving gel buffer (2.0 M Tris.HCl, pH 8.8)	3.000
40%Acrylamide (Acrylamide:Bisacrylamide, 37.5: 1)	4.500
Distilled water	5.505
0.5 M EDTA (pH 8.0)	0.015
8 M Urea	1.875
*Ammonium per sulphate (10 %)	0.075
TEMED	0.030

TEMED is *N,N,N′,N′*-Tetramethylethylenediamine. *Ammonium persulphate was freshly prepared before use and added last

Table 2.9 Urea-PAGE stacking gel casting recipes

Ingredient	Volume (mL)
Stacking gel buffer (0.5 M Tris.HCl, pH 6.8)	1.000
40%Acrylamide (Acrylamide:Bisacrylamide, 37.5: 1)	0.300
Distilled water	2.100
0.5 M EDTA (pH 8.0)	0.010
8 M Urea	0.500
*Ammonium per sulphate (10 %)	0.075
TEMED	0.015

TEMED is *N,N,N′,N′*-Tetramethylethylenediamine. *Ammonium persulphate was freshly prepared before use and added last

2.8.4 Optimisation of gene expression

Gene expression was performed after inserting recombinant plasmids into an expression host. This was achieved by transforming chemically competent *E. coli*

BL21 (DE3) with each of the recombinant plasmids. After transformation, mixtures were plated on LB agar (supplemented with ampicillin) and incubated at 37°C overnight. Antibody-resistant colonies (transformants) were selected and grown in 5 mL LB broth (supplemented with ampicillin) overnight at 37°C with shaking at 200 rpm and 1 mL portions were sub-cultured in 50 mL LB broth (supplemented with ampicillin). The 50 mL culture was grown at 37°C and 150 rpm to an OD_{600} of 0.5 – 0.7. Isopropyl β-D-thiogalactopyranoside (IPTG) was added to the cultures at different concentrations and incubated at different times before harvesting as shown in table 2.10.

Table 2.10 Conditions for gene expression

Flask	IPTG (mM)	Temperature (°C)	Time
1	0	37	4 hrs
2	0.1	37	4 hrs
3	0.5	37	4 hrs
4	1.0	37	4 hrs
5	1.0	16	Overnight

2.8.5 Large scale gene expression

In this case, single antibody resistant colonies (transformants) were selected and grown in 10 mL LB broth (supplemented with either ampicillin or kanamycin) overnight at 37°C with shaking at 200 rpm. This overnight preculture was used to inoculate one litre of terrific broth (containing appropriate antibiotic) and grown to an OD_{600} of 0.5 – 0.7 at 180 rpm, 37°C. Protein expression was induced by the addition of IPTG and incubated at the temperature that gave maximum protein expression following optimisation.

2.8.6 Production of lysates

After induction, the cells were harvested by centrifuging at 2880 x *g* for 10 min at 4°C. Pelleted cells were re-suspended to 10 mL/g in 0.02 M Tris.HCl (pH 7.4, containing 0.5 M NaCl, 0.03 M Imidazole) and lysed by exposing re-suspended cells (in 5 mL aliquots) to 6 cycles of 10 s sonication/ 10 s cooling with amplitude setting at 10 μ using Soniprep 150. Lysate was clarified by centrifuging at 27,000 x *g* for 30 min at 4°C.

2.8.7 Protein purification

Recombinant protein was purified from the clarified lysate *via* a two-step process – Affinity chromatography followed by ion-exchange.

.2.8.7.1 Affinity chromatography

The clarified lysate was subjected to Immobilised Metal Affinity Chromatography (IMAC) as the first step in the purification of the recombinant proteins. A Sepharose chelating fast flow resin column (Pharmacia Biotech) fixed to an automated AKTA purifier Fast Protein Liquid Chromatography (FPLC) system (Amersham Pharmacia) was used for IMAC. The Sepharose chelating fast flow resin was charged with Ni^{2+} by passing 0.1 M $NiSO_4$ through the column at a flow rate of 5 mL/min (for 6 min) under the control of the FPLC pump. The column was equilibrated with 200 mL of 0.02 M Tris.HCl (pH 7.4) containing 0.03 M Imidazole, 0.5 M NaCl (otherwise known as equilibration buffer) at a flow rate of 5 mL/min (for 20 min). The sample (*i.e.* clarified lysate) was loaded onto the column at a flow rate of 2 mL/min and washed with the equilibration buffer at a flow rate of 4 ml/min for 25 min. Elution of bound protein was effected *via* an imidazole linear gradient (From 0.03 M – 0.5 M)

with 0.02 M Tris.HCl (containing 0.5 M Imidazole, 0.5 M NaCl, pH 7.4) (otherwise known as elution buffer) using equilibration buffer as the diluent at a flow rate of 4 mL/min. Eluate was collected in 4 mL fractions. The composition of each fraction was analysed on SDS-PAGE to identify those containing protein of interest.

.2.8.7.2 Anion exchange chromatography

After SDS-PAGE of fractions following IMAC, fractions suspected to contain the protein of interest were pooled together and concentrated to about 0.5 mL by centrifuging at 2880 x *g*, 4ºC using a 10 kDa Vivaspin 20 Centrifugal Concentrator (Sartorius Stedim Biotech, France). This process was repeated three times with 20 mL of ice-cold 0.02 M Tris.HCl (pH 8.0) to remove imidazole and NaCl. After the final concentration the sample was diluted to a final volume of 20 mL in the same buffer and subjected to anion exchange chromatography. In this case, a 5 mL bed volume HiTrap Q HP (Hitrap ion exchange column, GE Healthcare) was attached to the FPLC system instead of the Sepharose chelating fast flow resin column. The column was equilibrated with 0.02 M Tris.HCl (pH 8.0) at a flow rate of 5 mL/min, sample applied at a flow rate of 2 mL/min and eluted with 0.02 M Tris.HCl, 1 M NaCl (pH 8.0) at a flow rate of 4 mL/min in a linear gradient using 0.02 M Tris.HCl (pH 8.0) as the diluent. Each fraction was analysed by SDS-PAGE.

2.8.8 Western Blotting

After SDS-PAGE, gels were removed from electrophoresis plates and soaked in transfer buffer (Appendix D) for 15 min. Nitrocellulose membranes and pieces of blotting paper (cut a little larger than the size of the gel) were also soaked separately in transfer buffer for 15 min. The base plate of the Transblot (Semi dry) apparatus

was cleaned and dampened with transfer buffer. Two soaked blotting papers were placed on the base plate followed by the nitrocellulose membrane and the gel. Two other soaked blotting papers were finally placed on the gel ensuring that no air bubbles were caught beneath the layers. The upper plate of the blotting apparatus was also cleaned and dampened with transfer buffer and the instrument was finally assembled. A current of 0.6 A was applied across the instrument for a single blot, 10 V for 90 min. Afterwards, the apparatus was dissembled and the nitrocellulose membrane was removed and stained with Ponceau S for 10 min to locate sample bands and standards. The membrane was washed with water after marking the standard bands with a pencil. The membrane was incubated in blocking solution (Appendix D) overnight at room temperature with gentle shaking. Membrane was then washed with phosphate-buffered saline-Tween 20 [(PBST) (Appendix D)] once and incubated with primary antibody (Table 2.11) for 90 min. The membrane was washed three times with PBST and incubated with secondary antibody (Table 2.12). The membrane was washed three times with PBST and developed with 5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium (BCIP/NBT conjugate). After colour development, the membrane was washed with water, dried and image visualised and captured using a Bio-Rad Gel Doc system with Quantity One software.

Table 2.11 Primary antibodies used for Western blotting

Primary/ labelling antibody	Dilution*	Analyte proteins
Monoclonal Anti-VapA MAb10G5	1/ 10,000	VapA
Hyperimmune Horse Serum	1/ 2,000	Immunoreactive *R. equi* proteins
Monoclonal Polyhistidine	1/ 3,000	Hexahistidine tagged proteins

* Diluted in blocking buffer

Table 2.12 Secondary antibodies used for Western blotting

Secondary antibody conjugate (alkaline phosphatase)	Dilution*	Primary antibody detected
Anti-mouse IgG	1/20,000	Monoclonal Anti-VapA
Anti-horse IgG	1/20,000	Hyperimmune Horse Serum
Anti-mouse IgG	1/20,000	Monoclonal Polyhistidine

* Diluted in blocking buffer

2.8.9 Protein Crystallisation, data collection, structure solution and refinement

After anion exchange chromatography, fractions containing purified proteins were pooled and buffer exchanged into 0.01 M Tris.HCl, 0.01 M NaCl, pH 7.5) and concentrated. Concentrated protein solution was subsequently centrifuged at 13,000 x *g*, for 5 min at 4ºC to removed particulates and undissolved substances. Crystallisation was performed according to the hanging drop vapour diffusion method. Crystal screens were performed in 24-multi well plates. Crystallisation conditions were Crystal Screen 1, Crystal Screen 2, PEG/Ion Screen, SaltRx (Hampton research), and Clear Strategy Screen 1 and Clear Strategy Screen 2 (Molecular Dimensions). Each mother liquor (500 µL) was delivered to a well and drops of 1 µL containing protein (30 mg/mL) were mixed with either 1 µL or 2 µL of the liquor drawn from the well. A drop was also constituted which contained 2 µl of protein and 1 µL of mother liquor. Drops were mixed on a siliconised coverslip (which has been polished with a silk scarf). The rim of each well (containing various mother liquors) was gently topped with high vacuum grease (Dow Corning) and the coverslip (containing the drops) was inverted and sealed above the well. In order to obtain crystals of good quality, the conditions in which promising crystals or precipitates developed were optimised by varying the concentration of salt,

precipitant or by adjusting the pH of the buffer. Crystals (after development) were harvested by transferring to a cryoprotectant (mother liquor (in which the crystal grew) containing 30 % glycerol) using a rayon fibre loop and stored in liquid nitrogen. Xray diffraction data were collected at 100 K at the Diamond Light Source experimental stations 102. Data were integrated using XDS (Kabsch, 2010) and scaled/merged with Aimless (Winn *et al*., 2011). Structure was solved by molecular replacement with the program MOLREP (Vagin and Teplykov, 2010) using a family related structure. The resulting model was refined using maximum likelihood methods implemented in REFMAC5 (Murshudov *et al*., 2011) with 5 % of the total data excluded from the refinement for the purpose of R free calculations. This procedure was interspersed with manual corrections to the model using COOT (Emsley *et al*., 2010) in conjunction with $2F_o$-F_c and F_o-F_c electron density maps.

2.8.10 Phase separation of Rhodococcus equi membrane proteins

This was done *via* Triton X-114 phase partitioning as described by Tan *et al*. (1995) with modifications. *Rhodococcus equi* 103^+ and *Rhodococcus equi*Δ*KasA* were grown at 37°C to mid-stationary phase in 50 mL LB broth (about 60 hr) with constant shaking at 150 rpm. Cells were harvested by centrifugation and washed with 2 mL of 0.01 M Tris. HCl (pH 7.4) containing 0.15 M NaCl (otherwise known as Tris buffered saline, (TBS)). Washed cells were re-suspended in 2 % Triton X-114 in TBS with phenylmethylsulfonylflouride (PMSF) (1 mM final concentration) and shaking overnight at 4°C and 150 rpm (About 30 mg - 50 mg of wet cells per mL of 2 % TX-114). Insoluble material was removed by centrifugation at 27,000 x *g* at 4°C for 15 min. The supernatant was warmed to 37°C for 10 min for phase partitioning and centrifuged at 25,155 x *g* at 25°C for 15 min. The upper aqueous phase was

removed and Triton X-114 was added to a final concentration of 2 % (v/v) and re-extracted. One millilitre of TBS was added to the lower detergent phase, mixed and phase partitioned again. The aqueous phases and detergents phases were pooled separately and proteins precipitated by 10 volumes acetone overnight at -20°C. The proteins were pelleted by centrifugation at 27,000 x *g* for 30 min at 4°C. The supernatant was removed, the pellet dried in a speed-vac, and subsequently re-suspended in 100 mL of Tris.HCl (0.01 M, pH 7.4).

2.8.11 Detection and identification of VapA from integral protein extract

The detergent and aqueous phase extracts were later subjected to dot blotting (for the wild type) in order to detect the presence of VapA. Protein identification was done by mass spectrometry after in-gel trypsin digestion.

.2.8.11.1 Dot Blotting

Grids were drawn with a pencil on a nitrocellulose membrane to indicate the regions to spot. Two microlitres of either the detergent phase extract or the aqueous phase extracts were dotted on the centre of each grid and allowed to dry under room temperature. The membrane was soaked in blocking solution (Appendix D) for 60 min at room temperature with gentle shaking. The membrane was washed once and incubated with Monoclonal anti-VapA, i.e. Mab 10G5 (1:10,000 dilution in blocking solution)], washed and later incubated with anti-mouse IgG, alkaline phosphatase conjugate (1:20,000 dilution in blocking solution). Membrane was washed and developed as in western blotting. Hexahistidine tagged VapA (expressed in *E. coli* BL 21 (DE3)) was used as a positive control.

.2.8.11.2 Liquid chromatography/Tandem mass spectrometry (LC-MS/MS)

Ten microlitres of the detergent phase extract was subjected to SDS-PAGE, stained by Coomassie blue and destained overnight. Thereafter, the gel was subjected to in-gel trypsin digestion before mass spectrometry. Briefly, the gel was placed on a cleaned white board and protein spots (already stained corresponding to the molecular size of interest) were excised using a scalp blade and transferred into clean 1 mL siliconised low retention microcentrifuge tubes. Two hundred microlitres of ammonium bicarbonate (100 mM) and 120 µL of acetonitrile (ACN) were added to each tube and the colloidal Coomassie blue stain was removed by votex for 30 min. The liquid was discarded and the wash repeated thrice. The gel slices were dehydrated by the addition of 50 µL of ACN and incubated at room temperature for 5 min. The liquid was discarded and further dehydrated twice. The gel slices were dried in a vacuum concentrator for 15 min, and incubated on ice with 25 µL of 20 µg/mL trypsin solution (Appendix D) for 30 min. Each gel slice was covered with 30 µL of 50 µM ammonium bicarbonate and incubated overnight at 37°C. The reaction was stopped by the addition of 30 µL of 50 % (v/v) ACN/ 5 % trifluoroacetic acid (TFA) and shaken for 30 min. The liquid (containing digested peptides) was recovered and transferred into a 1 mL microcentrifuge tube. Further digested peptides were recovered by shaking the gel for 30 min with 60 µL of 83 % (v/v) ACN/ 0.2 % TFA. The extract was pooled with the previous liquid containing the digested peptides and freeze-dried. Dried samples were re-suspended in 30 µL of 5 % ACN and 8 µL were subjected to liquid chromatography/ Tandem mass spectrometry (LC-MS/MS) using a HPLC Ultimate 3000 system (Dionex) coupled to an HCT Ultra mass spectrometer (Bruker Daltonics). Raw chromatographic data

were processed and matched to non-redundant peptides in the NCBI protein database using Mascot.

2.8.12 Expression and analysis of pEKEx2-derived plasmids in Corynebacterium glutamicum

A single transformant (from section 2.7.9) was selected and grown in 10 mL of LB-2 % glucose (containing kanamycin) overnight with shaking at 150 rpm at 30°C. Culture was inoculated in 50 mL LB-2 % glucose at a 1:100 dilution and grown (under same conditions) to an optical density of 1 (at 600 nm). Cells were subjected to IPTG induction (1 mM) and grown for either 4 hrs or 24 hrs with shaking at 150 rpm at 30°C. Each culture (50 mL) was centrifuged at 2880 x *g* for 10 min at 4ºC. The pellet was sonicated and clarified [otherwise called clarified lysate (CLY)] while the external culture medium was recovered and further clarified by centrifuged at 27,000 x *g* for 15 min at 4°C and sterile filtered. Trichloroacetic acid (TCA) was added to the clarified external culture to a final concentration of 10 % (w/v), mixed by gentle shaken and incubated at 4°C overnight. Precipitated proteins were harvested by centrifugation (27,000 x *g*, 4°C, and 30 min), re-suspended and washed three times in cold methanol. Washed samples were later dried using a speed vac for 15 min and re-suspended in 0.02 M Tris.HCl (pH 7.4) containing 0.5 M NaCl (solution containing proteins otherwise called broth supernatant proteins, BSP). Both CLY and BSP were subjected to SDS-PAGE and western blot using monoclonal anti VapA (Mab 10G5) as primary antibody and anti-mouse IgG (alkaline phosphatase conjugate) as secondary antibody. Protein identification was determined via mass spectrometry as described in section 2.8.11.2.

2.8.13 Subcellular compartmentalisation of VapA in Corynebacterium glutamicum

Transformants (after electroporation of pEKEx2-derived plasmids into *C. glutamicum*) were grown in 2 L LB-2 % glucose (containing kanamycin) at 30°C with shaking at 200 rpm. Isopropyl β-D-1-thiogalactopyranoside (1 mM) was added after the culture grew to OD_{600} of 1 and incubated at 30°C with shaking at 150 rpm. After overnight incubation, cells were pelleted by centrifuging at 2880 x *g*, 4°C for 10 min. Cells were re-suspended in Tris buffer, lysed by sonication, and clarified (as described). The supernatant was removed and the debris was re-suspended in 12.5 mM Tris.HCl pH 6.8 containing 2 % SDS (otherwise known as SDS-Tris buffer) or 12.5 mM Tris. HCl pH 6.8 containing 6 M urea (otherwise known as Urea-Tris buffer), and stirred overnight at 4°C and centrifuged at 27,000 x *g* for 30 min at 4°C for clarification. The clarified lysate (CLY) and the external culture medium (after further clarification by centrifuging at 27,000 x *g* and ultrafiltration (0.2 μm)) were subjected to high speed centrifugation at 100,000 x *g* for 2 hrs at 4°C. The pellets were also re-suspended in either SDS-Tris buffer or Urea-Tris buffer and extracted as in the debris. Each fraction was subjected to Western Blotting using Monoclonal Anti VapA (Mab 10G5) as the primary antibody.

2.8.14 Improved protocol for compartmentalisation of VapA in Corynebacterium glutamicum

The subcellular localisation of VapA in *C. glutamicum* was further fine-tuned. The debris after clarification of the lysate was subjected to Percoll density gradient centrifugation. A 60 % Percoll (v/v) solution was prepared in phosphate-buffered saline (as the relaxation buffer). The debris was re-suspended in Percoll solution (in

relaxation buffer) and centrifuged at 27,000 x *g* for 1 hr. The buoyant fraction was withdrawn and diluted with PBS and centrifuged at 27,000 x *g* for 1 hr. This was repeated until a fine pellet was obtained (which was washed free of Percoll). The pellet (cell wall material) was re-suspended in SDS-Tris buffer and stirred overnight and extracted as described. The protein contents of this cell wall fraction, the membrane fraction and the culture supernatant were standardised and subjected to SDS-PAGE and western blot.

2.8.15 Analytical Ultracentrifugation

Sedimentation velocity experiment was performed using a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics. Protein samples were dialysed overnight into 20 mM Tris.HCl (pH 7.9), 50 mM NaCl and loaded (with or without metals) into cells with 2-channel Epon centre pieces and quartz windows. Data were recorded at 40,000 rpm, 20°C overnight and a total of 200 absorption scans (280 nm) were recorded for each sample. Data analysed was done using the SEDFIT software fitting a single frictional coefficient.

2.8.16 Interaction of Virulence Associated Proteins

The interaction of *Rhodococcus equi* virulence associated proteins was done *via* Surface Plasmon Resonance using a BIAcore X100 at 25°C. The surface of a CM5 censor chip (Carboxymethylated dextran covalently attached to a gold surface) was activated with a 1:1 mixture of 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) at a flow rate of 5 µl/min for 1 min. VapA, dialysed in 10 mM acetate buffer (pH 3.5), 5 mM NaCl overnight, was diluted to 42 µg/mL and immobilised on the activated sensor chip at 5 µl/min for 1 min. When

sufficient VapA was immobilised, reactive groups remaining on the activated CM5 surface were blocked with 1 M ethanolamine-HCl pH 8.5 at a flow rate of 5 µl/min for 1 min. Other Vaps (different concentrations in 10 mM HEPES, pH 7.4, containing 150 mM NaCl, EDTA 3 mM, T20 0.005%), were passed over the VapA-coated and control coated cells at a flow rate of 30 µl/min for 3 min. After every cycle, the chip was regenerated by passing through 10 mM glycine.HCl (pH 2.0) for 1 min at a flow rate of 30 µL/min.

2.8.17 Modelling of Virulence Associated Proteins

Where applicable the 3D structures of *R. equi* virulence associated proteins were modelled using Modeller 9.13 (Sali, 1995). The reliability of the structures was validated by COOT (Emsley *et al*., 2010) and ProSA-web (Wiederstein and Sippl, 2007) using the pdb files generated by the homology modelling.

2.9 Lipid Analysis

2.9.1 Extraction and analysis of lipids from phase separated complexes

.2.9.1.1 Biphasic extraction of free lipids

Stock containing phase–separated proteins (from section 2.8.10) were re-suspended in ice-cold trichloroacetic acid (final concentration of 10 % TCA) and incubated on ice for 3 hrs. These samples were centrifuged at 27,000 x *g* for 30 min at 4°C and lyophilised overnight after discarding the supernatant. The residue was re-suspended in 2 mL methanol:0.3 % NaCl (100:10 v/v) and transferred to a 12 mL clear screw capped borosilicate glass tube. Two millilitres of petroleum ether (High boiling point) was added and mixed thoroughly using a rotator for 30 min. The solution was then centrifuged at 720 x *g* for 5 min at 4°C. The upper phase (lipid extract) was

removed into a separate tube and the lower phase washed again. The lower phase (delipidated) was lyophilised overnight while the non-aqueous lipid phase was dried using a heating block.

.2.9.1.2 Thin layer chromatography

The residue (after drying of the upper non-aqueous phase) was re-suspended in 200 µL of chloroform:methanol (2:1) and 10 µL was spotted on a pre-coated TLC-sheet (Alugram^{R} SIL G/UV$_{254}$). It was developed in chloroform:methanol:water (60:16:2) mobile phase. After development, bands were revealed by charring using a heat gun after staining with 5 % ethanolic molybdophosphoric acid.

.2.9.1.3 Extraction and analysis of protein-bound lipids

The residue after lyophilisation of the delipidated extract was re-suspended in 1 mL of 5 % aqueous tetrabutylammonium hydroxide and heated overnight at 100°C using a heating block and allowed to cool. Distilled water (2 mL), dichloromethane (4 mL), and iodomethane (250 µL) were added and mixed for 1 hr using a rotator. The solution was centrifuged for 5 mins at 720 x *g* at room temperature and the upper phase discarded. The solution was washed by the addition of 4 mL distilled water, mixed for 30 min and centrifuged for 5 min at 720 x *g*. The upper phase was discarded and the solution washed twice using 4 mL of distilled water. The solution was dried overnight using a heating block and re-suspended in 4 mL of diethyl ether, mixed and centrifuged for 5 min under room temperature at 720 x *g*. The supernatant was transferred to a clean tube and dried at 50°C. The final residue was dissolved in 100 µL of dichloromethane and subjected to thin layer chromatography using

petroleum ether: acetone (95:5) development and visualised by charring after staining in 5 % ethanolic phosphomolybdic acid.

2.9.2 Analysis of protein-bound lipids from Corynebacterium glutamicum transformants

Proteins were precipitated from the CLY and the BSP (obtained from section 2.8.12) using trichloroacetic acid. Samples were delipidated, protein-bound lipids extracted and subjected to TLC analysis as described previously.

2.9.3 Extraction of cell wall associated lipids

Cell wall associated lipids were extracted from whole cells (both from *Corynebacterium glutamicum* and *Rhodococcus equi*) according to the method of Kacem *et al.* (2004) with some modifications. Cells were grown in BHI broth overnight and harvested via centrifugation. The lipids were first extracted from wet cells (100 mg) with chloroform: methanol (1:1) overnight using a rotator at room temperature. It was centrifuged and the bacterial residue was re-extracted three times using chloroform:methanol (2:1). The lower organic phases were all pooled and evaporated to dryness in a heating block at 50°C. The crude extract was biphasically-partitioned using a mixture of chloroform:methanol:water (8:4:2) and the lower organic phase recovered and dried. Residue was finally re-suspended in 200 µL of chloroform:methanol (2:1) and 10 µL was subjected to thin layer chromatography developed in chloroform:methanol:water (60:16:2) and bands visualised by charring after staining in 5 % ethanolic phosphomolybdic acid. The rest was dried and stored at room temperature for further use.

2.9.4 Matrix-assisted laser desorption/ ionisation-time of flight mass spectrometry (MALDI-TOF MS)

Protein-bound lipids (derivatised as methyl esters) both from *Rhodococcus equi* 103S and *Rhodococcus equiΔkasA* (following Triton X-114 phase separation) were dried by evaporation at 50°C using a heating block. Dried samples were re-suspended in a 1:1 mixture of acetonitrile (ACN) and α-Cyano-4-hydroxycinnamic acid (HCCA, which served as the absorbing matrix) preparation. Each sample (1.5 µL) was loaded on a target plate and subjected to MALDI-TOF mass spectrometry using an Ultraflex II TOF-TOF (Bruker Daltonics). Where the sample was too viscous, it was diluted with ACN and 1 µL of the solution was mixed with HCCA and loaded on the target plate.

2.10 Other methods

2.10.1 Mycoloyltransferase assay

Cell wall associated lipids (from *Corynebacterium glutamicum* and *Rhodococcus equi*) were dried in a heating block and re-suspended in 0.02 M Tris. HCl pH 7.5, 0.01 M Dithiothreitol and sonicated. Recombinant VapA and corynomycolyl transferase (Cmt1) were dialysed against 2 x 2 litres of 0.01 M Tris.HCl pH 7.5 (containing 0.01 M NaCl) for 24 hrs. Ten microliters of VapA (at 1 mg/mL) was mixed with 5µL of each lipid preparation. The reaction was initiated by the addition of 10 µL of corynomycolyl transferase (at 1 mg/mL) and incubated for 45 min. Thereafter, reaction mixtures were subjected to Urea-PAGE and visualised by Coomassie blue staining as reported. The Cmt1 assay was also performed using a radioactive substrate. In this case, ^{14}C-labelled free lipids extracted from

Mycobacterium bovis BCG were used for the assay. After assay, reaction mixtures were subjected to Urea-PAGE. Gels were dried and exposed and autoradiograms were produced by 12 hrs exposure to Kodak BioMax MR film to reveal ^{14}C labelled lipids (bound to proteins).

CHAPTER THREE

3 Cloning, purification and initial characterisation of Vap proteins

3.1 Introduction

The Vap proteins of *R. equi* display significant sequence similarity (especially at their carboxyl regions) and very few homologues have been found in any other organism (Letek *et al*., 2008). Apart from VapF, they also possess predicted signal sequences of secreted proteins; with well-defined *N*-terminal, hydrophobic, and *C*-terminal regions and potential signal peptidase cleavage sites (Takai *et al*., 2000; Byrne *et al*., 2001) (Figure 1.17). VapA is retained at the cell surface and apparently undergoes an unusual lipid modification. At the outset of this study, neither the mechanism *via* which Vaps enhance virulence nor their molecular structures were defined. This chapter describes the cloning, expression, purification of the Vap proteins carried out in order to facilitate the studies in subsequent chapters that focus on their structural characterisation and the exploration of potential homologous and heterologous interactions of Vaps. VapF was not considered here since its gene has a 5' truncation and two frame-shift mutations at the 3' end; expression would result in non-secreted and a presumed non-functional protein (Takai *et al*., 2000: Letek *et al*., 2008).

The coding sequences for the mature secreted proteins of *vapACDEGH* were cloned in order to produce high yield soluble protein. The cloning vectors pET23a or pET28a were used as they represent high copy number plasmids with strong T7 promoters. These plasmids facilitate hexahistidine tagging at the *C*-terminus (for pET23a) or the *N*-terminus (for pET28a) of the recombinant protein product which

can be exploited for purification *via* metal chelate affinity chromatography in a single step and immuno-detection with anti-polyhistidine.

3.2 Amplification of Vap genes

All the sequence data for ORFs that code for the Vaps under consideration were obtained from the National Centre for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov, accession codes in Table 3.1) and the amino acid sequences were analysed for the presence of signal peptide using the SignalP server (http://www.cbs.dtu.dk/services/SignalP) (Nielsen *et al.,* 1997; Petersen *et al*., 2011) using default settings for Gram-positive bacteria. The possession of the predicted signal sequence in the protein may direct the proteins for secretion when the genes are expressed in *E. coli* thus could significantly limit protein yield and function. Hence primers were designed in order to amplify the regions that code for the putative mature protein for cloning into pET23a and pET28a vectors. The sizes of the genes and the sequences that encode the mature proteins are shown in table 3.1.

Table 3.1 Summary of *vap* family properties

vap gene (Accession code)	Length / bp		Mature *N* -terminus
	gene	amplicon	
A (NP_002576.1)	570	477	T32
C (AF116907.2)	525	438	N30
D (AF118814.1)	495	405	Q31
E (AF116907.2)	621	483	R47
G (AF116907.2)	519	441	E27
H (AF116907.2)	564	471	A32

Primer pairs were designed such that each oligonucleotide had a minimum of 26 bases and a maximum of 32 bases in length and the difference between their individual melting temperatures was limited to a maximum of 2ºC. For the

amplification of *vapA* and *vapC*, the restriction sites for *Nde*I and *Not*I were added to the forward and reverse primers respectively. For the amplification of the other *vaps*, the sites for *Nde*I and *EcoR*I were added to the forward and reverse primers respectively.

The regions which encode the Vap proteins were amplified *via* polymerase chain reaction using Vent DNA polymerase as described (Section 2.7.4) and analysed *via* agarose gel electrophoresis (Figure 3.1). Initial attempts that used *R. equi* total DNA as a template were without success. Various protocols were deployed that analysed the use of Mg^{2+} and DMSO as additives and the variation of primer annealing temperatures through a gradient PCR approach. After rigorous testing, the quality of the DNA was validated in its successful use as a template for chromosomal genes used in other studies in this laboratory (REQ28150 and REQ09720, data not shown). At this stage, two possibilities were considered; firstly that the strain in use had been cured of its virulence plasmid and secondly that the plasmid did not extract well in the protocol used to purify the genomic DNA. The former scenario was excluded as VapA could be detected in extracts of these cells using the monoclonal antibody MabG105 (see figure 6.1) and VapA peptides could be detected in Triton X-114 extracts (see section 6.2.1). The second scenario seemed plausible. In order to ameliorate plasmid loss in extraction, attempts were made to amplify the *vap* genes using a small sample of freshly-cultured biomass as a crude DNA source, this too proved unsuccessful.

In order to make timely progress, a pragmatic solution was sought. Cosmid p10.1, which carries *vapA*, *vapD* and *vapC* was used as a template and other *vap* sequences

were synthesised using a service at Eurofins MWG. These were cloned in pET23a consistent with our original cloning strategy.

These alternative templates delivered high quality amplicons (Figure 3.1) with single bands albeit with some primer dimer evident. The migration of observed DNA bands was consistent with the expected amplicon sizes (Table 3.1); bands were excised from the gel and their DNA was purified from the agarose as discussed in section 2.7.3.

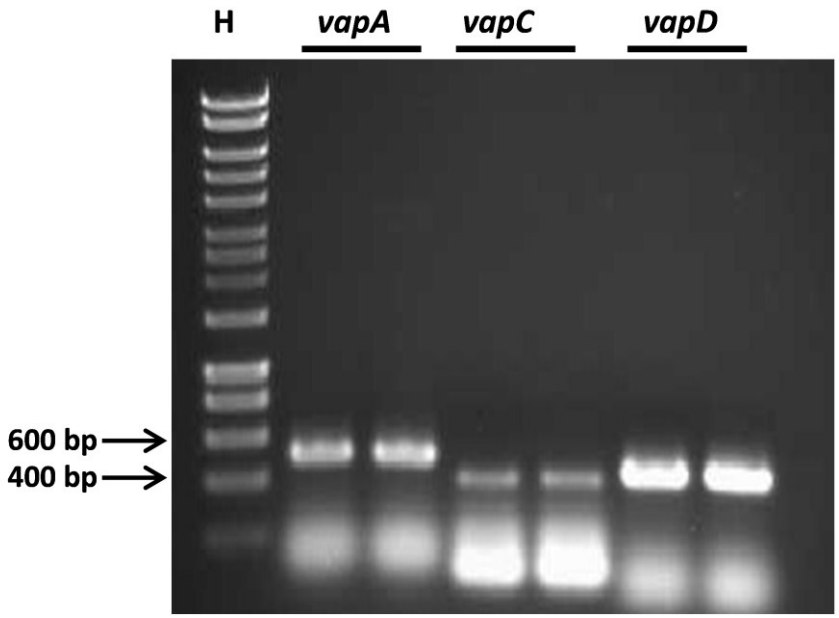


Figure 3.1 Preparation of coding sequences for *C*-terminally hexahistidine-tagged recombinant *vapA*, *vapC* and *vapD* by PCR amplification. Cosmid p10.1 was used as the DNA template. Lane H carries a DNA size standard (Hyperladder I).

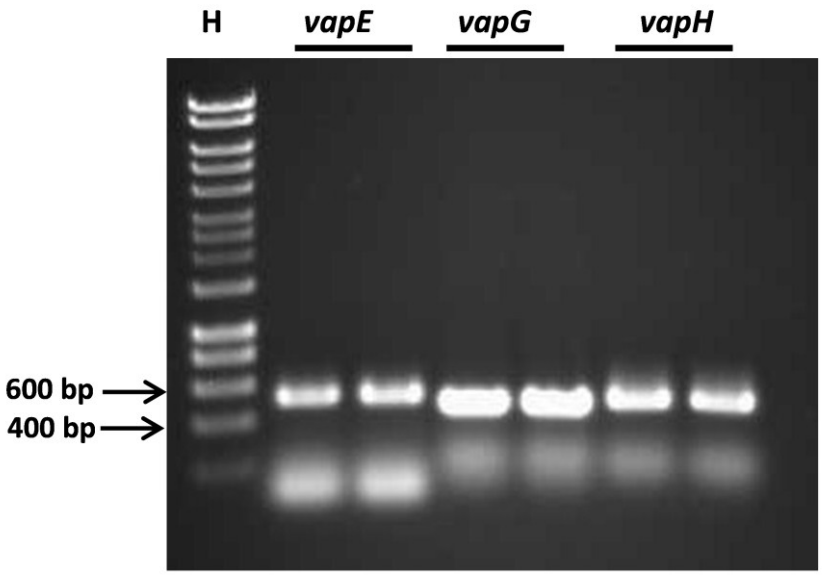


Figure 3.2 PCR amplification of *vapE*, *vapG* and *vapH* from validated pET23 clones to form pET28 variants.
Lane H, size standard (Hyperladder I)

3.3 Cloning of amplified products to vector and analysis

The purified amplicons were prepared for cloning by restriction digestion, and were ligated with the similarly-treated pET vectors (section 2.7.6). The ligation mixtures were used to transform competent *E. coli* Top 10 and transformants were selected in solid culture with appropriate antibiotics. Plasmid DNA was extracted from selected transformants after small-scale broth culture and were screened *via* restriction profiling (with the enzymes used for cloning) analysed by agarose gel electrophoresis (figures 3.3 and 3.4).

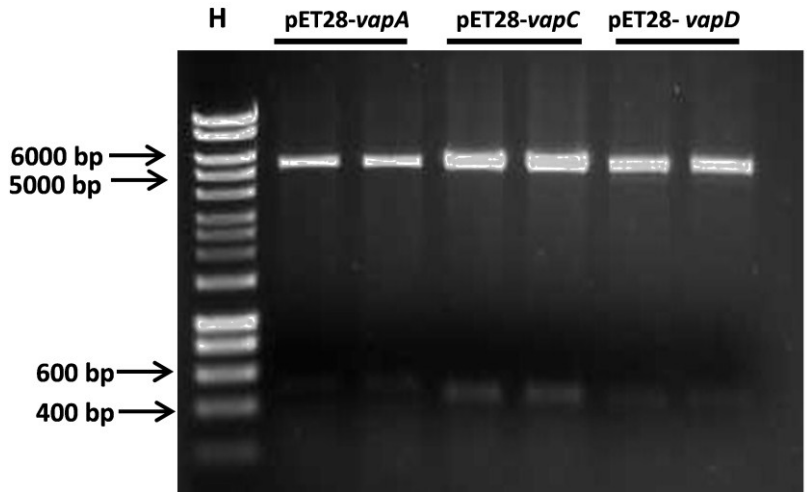


Figure 3.3 Screening for recombinant plasmids pET28-*vapA*, pET28- *vapC*, and pET28-*vapD*.
Plasmid DNA recovered from transformants was digested using restriction enzymes used in cloning (*vapA*& *vapC*, *Nde*I and *Not*I; *vapD* *Nde*I and *EcoR*I). Lane H carries a DNA size standard (Hyperladder I).

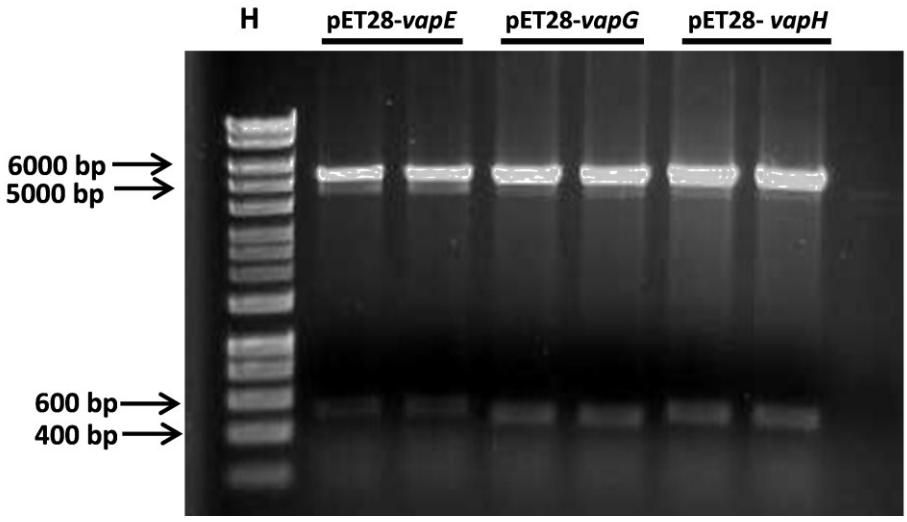


Figure 3.4 Screening for recombinant plasmids pET28-*vapE*, pET28-*vapG*, and pET28-*vapH*.
Plasmid DNA recovered from transformants was digested using restriction enzymes used in cloning, *Nde*I and *EcoR*I. Lane H carries a DNA size standard (Hyperladder I).

The banding patterns generated in this screening were compatible with expectations. The fluorescing band that migrates between the 5000 and 6000 bp markers is consistent with pET 28a, a high copy number plasmid of 5369 bp in length. DNA from supposed *vap* clones all generated bands that migrated consistent with expectations as defined in Table 3.1, which strongly implied that the *vaps* were all effectively cloned in the expression vector. Nucleotide sequencing confirmed the identity of the fragment, that it was intact and that the predicted gene product bore no mutations.

3.4 Recombinant protein production and purification

The recombinant plasmids were used to transform *E. coli* BL21 (DE3). Heterologous protein overproduction in *E. coli* depends on a number of factors especially temperature which has pronounced effect on protein folding and stability (Weickert *et al*., 1996). In order to determine the optimum condition for the production of soluble recombinant protein, cultures were grown to mid-log phase and induced with 1mM IPTG after cooling to an induction temperature between 16°C and 37°C. Electrophoretic analysis of clarified lysates (data not shown) prepared from the recombinant cells showed that synthesis of most of these recombinant proteins was induced under all conditions but was more pronounced when the recombinant gene expression was induced using 1mM IPTG induction with incubation at 16°C overnight. This practice was adopted as standard for large scale production of Vaps in *E. coli*. High yields of VapC were never achieved. In order to make progress in functional and structural studies, experiments proceeded without this protein.

The process was scaled up to 2 L cultures in terrific broth. Cells were harvested, lysed in 0.02 M Tris.HCl; 0.5 M NaCl; 0.03 M Imidazole, pH 7.4. Lysates were clarified (as described in section 2.8.6) and subjected to affinity chromatography. Fractions were eluted (monitored at 280 nm) in a linear gradient of imidazole concentration terminating at either 0.5 M or 1 M Imidazole as appropriate. Fractions were subjected to SDS-PAGE to analyse the homogeneity of those that contained recombinant protein. Peak fractions were pooled, concentrated and adjusted to pH 8.2. The solution was finally subjected to anion exchange chromatography using 0.02 M Tris.HCl (pH 8.2) as equilibration buffer and 0.02 M Tris.HCl (containing 1 M NaCl, pH 8.2) as elution buffer. Elution was monitored at 280 nm and peak fractions were subjected to SDS-PAGE. The SDS–PAGE gel images for VapA and VapG at each stage of purification are shown in figures 3.5 - 3.8. The images for VapD, VapE and VapH are shown in appendix F.

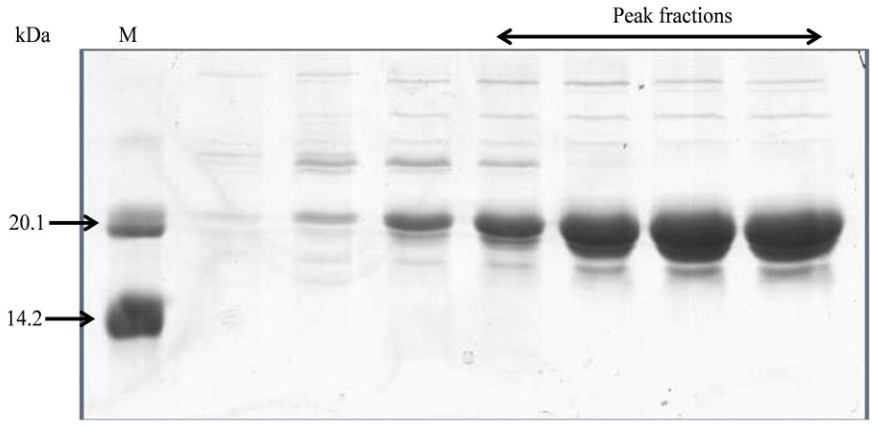


Figure 3.5 SDS-PAGE of recombinant VapA after immobilised metal affinity chromatography of clarifed lysate.
Successive fractions approaching peak elution as indicated by A_{280nm} were loaded. 'M' indicates molecular weight markers (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

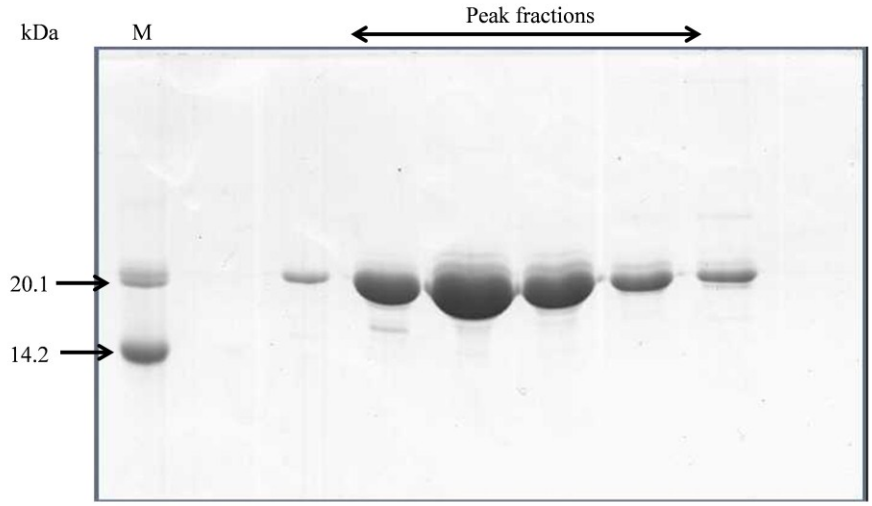


Figure 3.6 SDS-PAGE of recombinant VapA following anion exchange chromatography.
Successive fractions approaching peak elution as indicated by A_{280nm} were loaded. 'M' indicates molecular weight markers (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

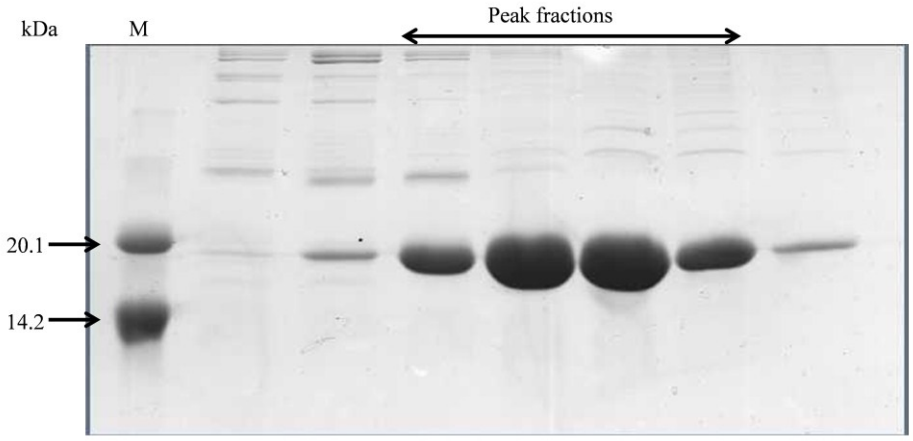


Figure 3.7 SDS-PAGE of recombinant VapG after immobilised metal affinity chromatography of clarified lysate.
Successive fractions approaching peak elution as indicated by A_{280nm} were loaded. 'M' indicates molecular weight markers (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

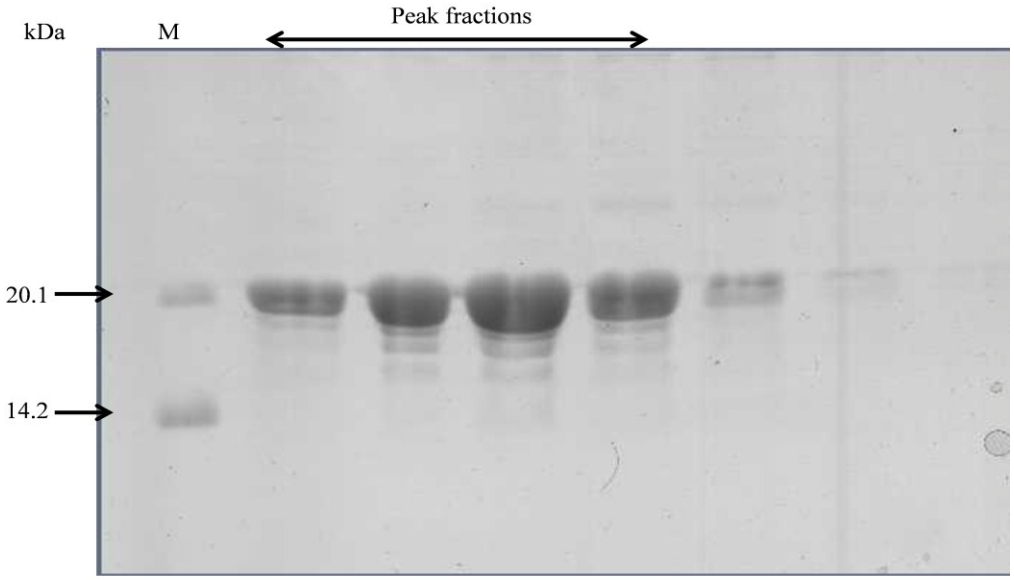


Figure 3.8 SDS-PAGE of recombinant VapG following anion exchange chromatography.
Successive fractions approaching peak elution as indicated by A_{280nm} were loaded. 'M' indicates molecular weight markers (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

During anion exchange chromatography, the protein eluted as a symmetrical peak and the purity of the proteins were revealed following SDS-PAGE. Despite some minor degradation, proteins were highly purified. Peak fractions were pooled and concentrated for further analysis.

3.5 Detection of Recombinant protein

Following affinity chromatography, a sample from a peak fraction for each protein was subjected to Western Blot on nitrocellulose. Blotted proteins were probed using various detection systems based on convalescent horse serum, monoclonal anti-polyhistidine and monoclonal anti-VapA (Mab10G5) with complementary secondary antibody conjugates (see section 2.8.8 and Figures 3.9 - 3.14).

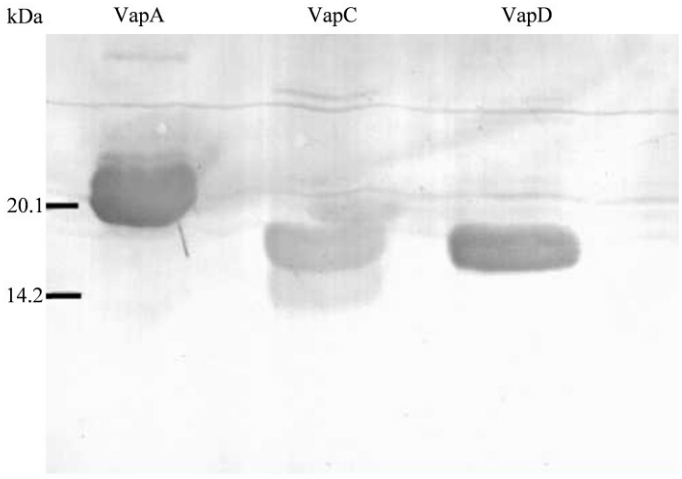


Figure 3.9 Immunodetection of recombinant VapA, VapC and VapD using convalescent horse serum.

Binding of antibodies from polyclonal convalescent serum to recombinant Vaps was monitored using anti-horse alkaline phosphatase conjugate. Migration of size standards was monitored using Ponceau S and is indicated at the left margin (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean)

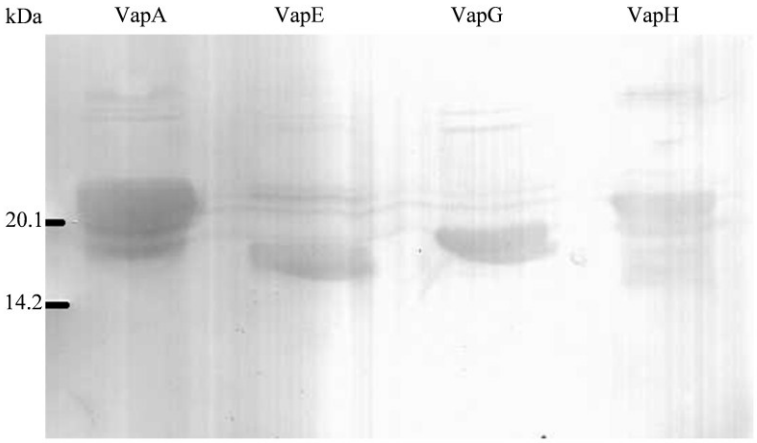


Figure 3.10 Immunodetection of recombinant VapA, VapE, VapG and VapH using convalescent horse serum.

Binding of antibodies from polyclonal convalescent serum to recombinant Vaps was monitored using anti-horse alkaline phosphatase conjugate. Migration of size standards was monitored using Ponceau S and is indicated at the left margin (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean)

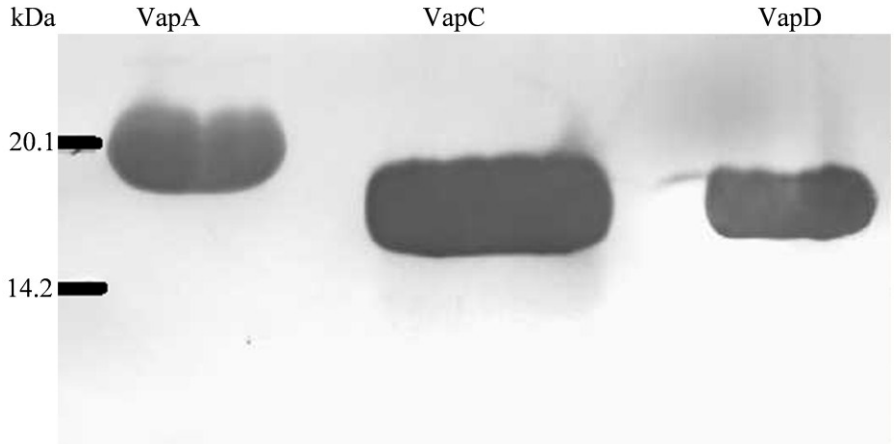


Figure 3.11 Immunodetection of recombinant VapA, VapC and VapD using monoclonal antipolyhistidine.

Binding to recombinant Vaps was monitored using goat-anti-mouse alkaline phosphatase conjugate. Migration of size standards was monitored using Ponceau S and is indicated at the left margin (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean)

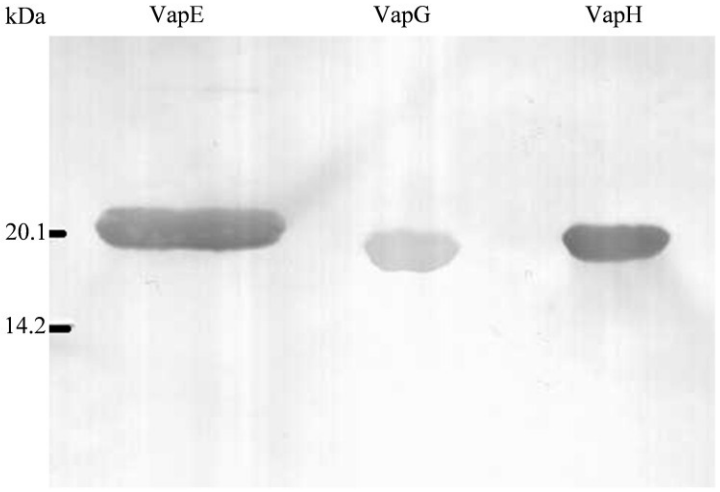


Figure 3.12 Immunodetection of recombinant VapE, VapG and VapH using monoclonal antipolyhistidine.

Binding to recombinant Vaps was monitored using goat-anti-mouse alkaline phosphatase conjugate. Migration of size standards was monitored using Ponceau S and is indicated at the left margin (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

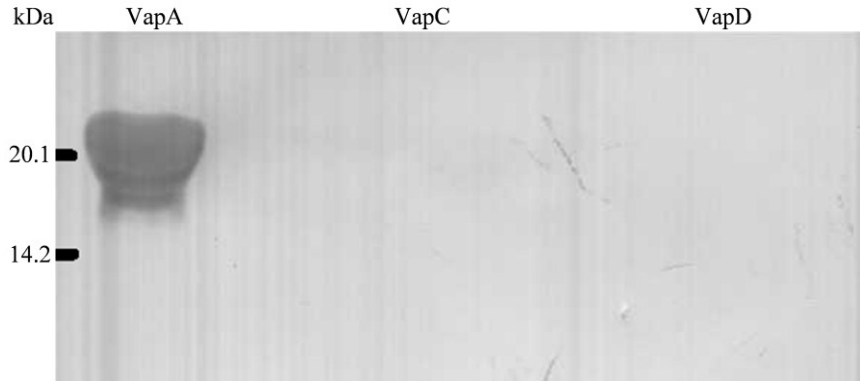


Figure 3.13 Immunodetection of recombinant VapA, VapC and VapD using monoclonal anti-VapA (Mab10 G5).
Binding to recombinant Vaps was monitored using goat-anti-mouse alkaline phosphatase conjugate. Migration of size standards was monitored using Ponceau S and is indicated at the left margin (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

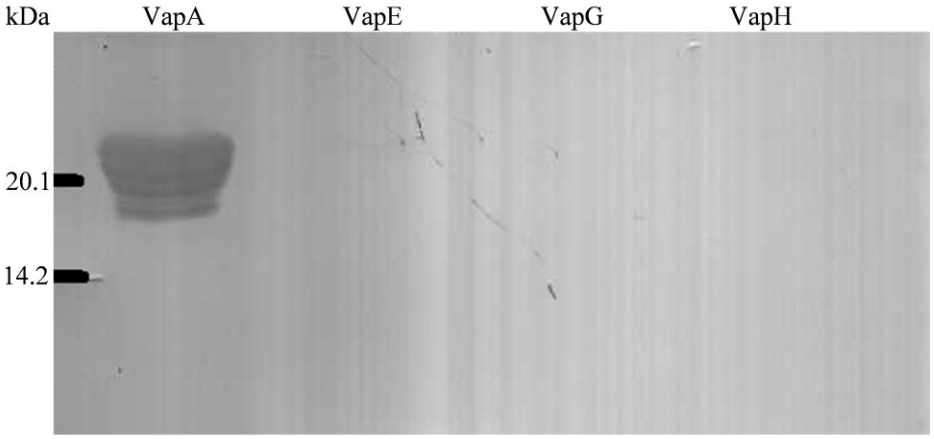


Figure 3.14 Immunodetection of recombinant VapA, VapE, VapG and VapH using monoclonal antiVapA (Mab10 G5).
Binding to recombinant Vaps was monitored using goat-anti-mouse alkaline phosphatase conjugate. Migration of size standards was monitored using Ponceau S and is indicated at the left margin (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

The banding profiles revealed that all of the recombinant proteins were detected with the hyperimmune horse serum, which strongly suggests the presence of antibodies to each Vap in the convalescent adult horse. This indicates that the native proteins exhibit significant immunogenicity in rhodococcal infection. All proteins reacted significantly with monoclonal anti-polyhistidine, which confirmed that the proteins were hexahistidine tagged. However, only VapA reacted with Mab 10G5, validating its use as a detection method for VapA elsewhere in this study. Though these proteins exhibit some degree of homology (especially in the *C*-terminal region, Figure 1.17), it seems they do not exhibit cross reactivity. In all cases, the banding pattern was consistent with the expected molecular sizes of the proteins.

In summary VapA, VapD, VapE, VapG, and VapH were produced at high yield and were purified to near homogeneity with a two-step chromatography protocol.

CHAPTER FOUR

4 Vap Structural models: crystallography and homology modelling approaches

4.1 Protein crystallisation

Each purified protein was exchanged to 0.01M Tris.HCl pH 7.5, 0.01 M NaCl, concentrated to 30 mg/mL and screened for the production of crystals using the vapour diffusion method as described in section 2.8.9. Only VapG produced promising precipitates in three conditions of the PEG/ION screen which were condition 8 (0.2 M potassium chloride, 20 % PEG 3350), condition 14 (0.2 M lithium nitrate, 20 % PEG 3350) and condition 15 (0.2 M potassium thiocyanate, 20 % PEG 3350). These conditions were optimised by varying either the concentration of the salt or PEG or both and screening performed with the protein. Good crystals developed between four and eleven days (Figure 4.1).

4.2 Structure solution and refinement for VapG

Structural analyses of crystals produced at Northumbria were undertaken in collaboration with the Wilkinson Group based at University of York. Preliminary diffraction of X-rays with these crystals identified that VapG crystallised in 0.2 M potassium chloride, 22 % PEG 3350 produced high quality data. Crystals were transported to the Diamond Light Source, the UK's national synchrotron facility for analysis on Beamline I02, a high throughput and highly automated beamline for optimised MAD and SAD experiments (wavelength 0.5Å - 2.5Å). The crystal exhibited a tetragonal space group with unit dimensions a = 44.55, b = 44.55, and c = 259.67 Å (Table 4.1).

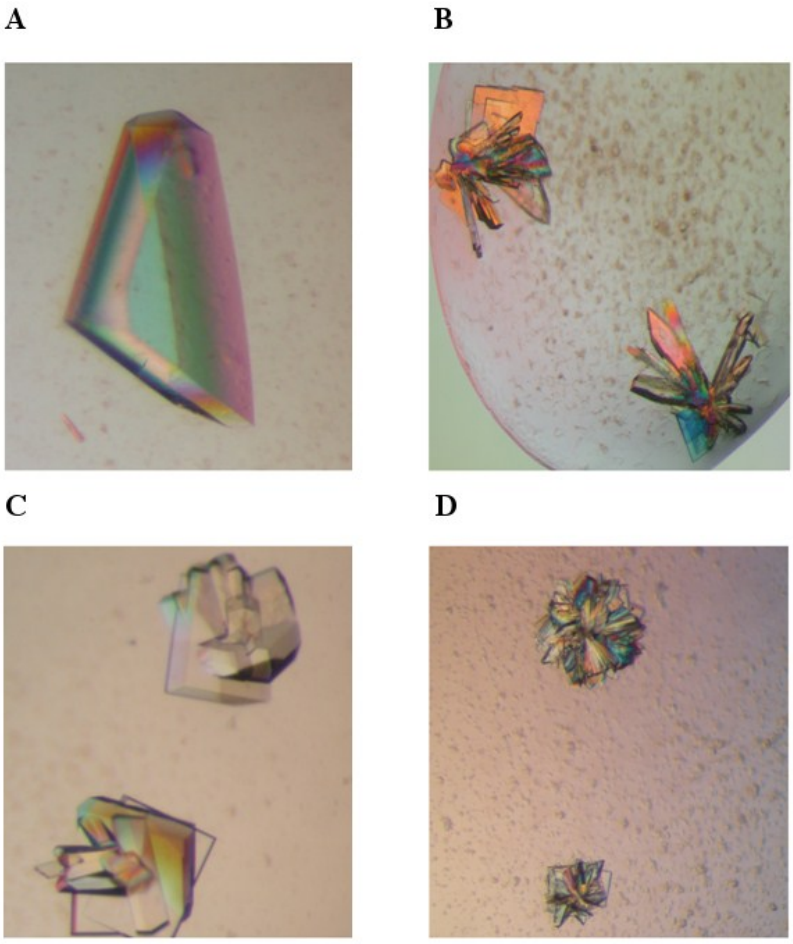


Figure 4.1 Crystals of VapG grown using the hanging drop vapour diffusion method.
Crystals developed in A; 0.2 M potassium chloride, 22 % PEG 3350 after four days, B; 0.3 M potassium chloride, 22 % PEG 3350 after 11 days; C; 0.2 M lithium nitrate, 22 % PEG 3350 after 6 days and D; 0.2 M lithium nitrate, 20 % PEG 3350 after 6 days.

The structure of VapG was solved by molecular replacement (as described in sections 2.8.9) using the newly solved coordinates of VapD (PDB code 4CSB, Whittingham *et al.*, 2014) as the starting model since the proteins share a 55 % sequence identity (Figure 4.2). Refinement statistics and structural information are listed in Table 4.1.

Figure 4.2 Amino acid similarity between VapD and VapG
Identity matrix was performed using the Clustal Omega 1.2.1 multiple sequence alignment program without the predicted signal sequences of the proteins. The proteins share 55 % identity. The nature of the aminoacyl side-chains are indicted by residue colour; hydrophobic (red), polar (green), acidic (blue) and basic (pink). Conserved amino acid identity is indicated beneath the aligned sequences by a star, similarities are indicated by colons and full stops.

The refined model of VapG consisted of two crystallographically independent protein molecules (molecules 1 and 2) which are almost identical (rms Δ = 0.45 Å calculated on all C_α atoms) (Figure 4.3). During refinement, elongated peaks of positive difference electron density appeared in two independent locations in the asymmetric unit. These peaks, although, generally lacking detail, had the appearance of additional protein chains or peptides bound between the protein molecules: in one location three histidine side chains were successfully built and refined; in the other a lysine-leucine pair were similarly identified and included in the model. Both elements likely derive from the artificial *C*-terminus (PNSSSVDKLAAALEHHHHHH) used for protein tagging. Difference electron density in the vicinity of these side chains was less easy to interpret and so was modelled with connecting alanine side chains only.

Table 4.1 Crystallisation conditions, data collection and refinement statistics

Data Collection and Processing Statistics	
Diamond beamline / wavelength (Å)	I02 / 0.9795
$P4_12_12$ cell dimension (Å)	*a, b* = 44.55, *c* = 259.67
Resolution limits (Å)	31.50 – 1.80 (1.85 – 1.80)[a]
No. unique reflections	24979
Completeness (%)	98.1 (83.4)
Multiplicity	11.4 (7.2)
R_{merge}*	0.043 (0.674)
I/σI	31.4 (2.6)
Refinement Statistics	
R_{cryst}/R_{free}	0.2009/0.2362
r.m.s. Δ bond length (1-2) (Å)	0.021 (0.019)
r.m.s Δ angles (°)	1.893 (1.925)
r.m.s Δ chiral volumes ($Å^3$)	0.126 (0.200)
Average B ($Å^2$): Protein/waters and K+ ion	31.55/38.02
Ramachandran outliers (%)	0

[a]Highest resolution shell statistics given in parentheses. [b]Average geometric restraints given in parentheses. *R_{merge} = Σ|I - <I>| / ΣI

The only notable difference between molecules 1 and 2 of the asymmetric unit electron density maps was a strong metal ion peak located at the centre of a cluster of main chain and side chains oxygen donors from residues Asn107 – Asn114 of molecule 1 (Figure 4.7). No such coordination was seen in molecule 2. The coordination distances ranged from 2.7 Å to 3.1 Å, which is consistent with a potassium ion (Harding, 2002). Although other metal ion candidates (calcium and magnesium) were tested in refinement, only potassium refined well. Furthermore, X-ray absorption analysis of another single crystal revealed no other metal ions present apart from potassium.

In both molecules, the *N*-terminal residues (Met1-Ala41) are apparently disordered. It is unlikely that the *N*-terminus is degraded (and therefore absent) since the crystals of VapG grew over a period of days rather than weeks or months. This finding is consistent with those of others (VapD in Whittingham *et al.*, 2014; VapB in Geerds *et al*., 2014) suggesting the notion that in all *R. equi* Vap proteins, this part of the protein forms random coil. By contrast, the *C*-termini of both protein molecules are well ordered, even extending into the *C*-terminal His_6 affinity tag.

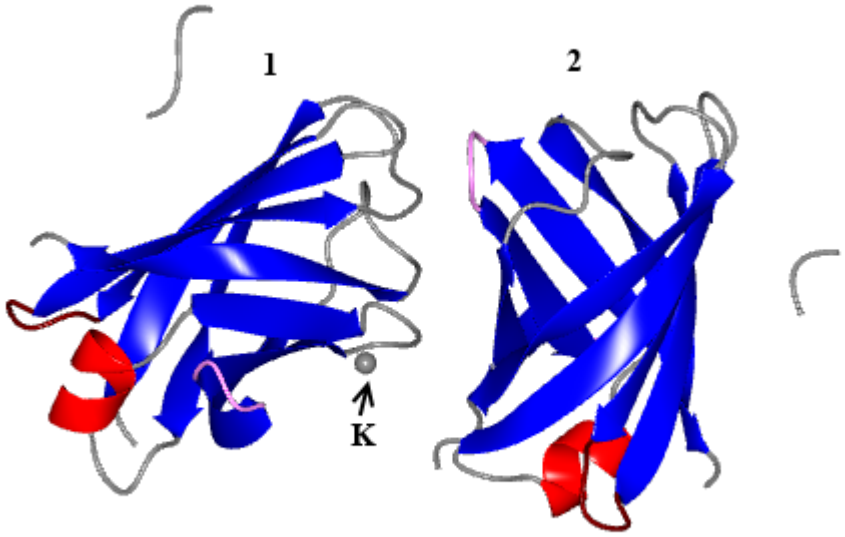


Figure 4.3 Structure of VapG showing two molecules in the asymmetric unit
Both structures are similar (rmsΔ = 0.45 Å) apart from that molecule 1 has a potassium binding site (The potassium ion is indicated by an arrow in molecule 1). Also modelled in the structure are the peptides KLAA (as a fragment in molecule 1), and HHH (as a fragment in molecule 2). Figure was created using CCP4mg (McNicholas *et al*., 2011)

The protein fold of VapG consists of a compact, 8-stranded β-barrel, which is elliptical in cross section with a single helix (Figure 4.3). The β-strand ordering in

the barrel is non-consecutive (β1- β2- β3- β8- β5- β6- β7- β4) which requires a strand crossover at both ends, hence the barrel adopts a closed form. At one end of the barrel (subsequently termed the top), the turns between strands are very short, providing a smooth, rounded surface, while at the other end (bottom), there are longer inter-strand regions which protrude in the form of a 9-residue α-helix (β4- β5) and two flanking loops (between β2- β3 and β6- β7) (Figure 4.4).

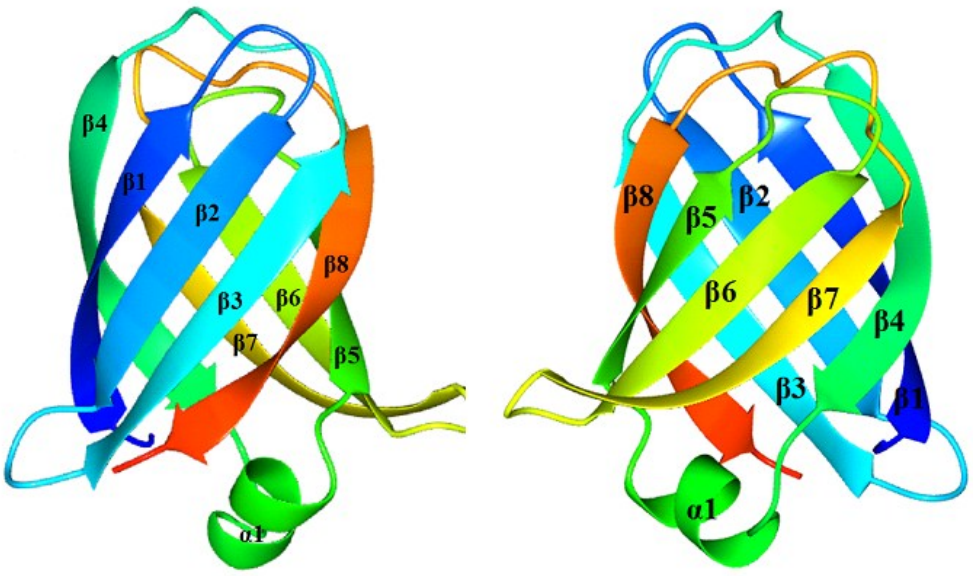


Figure 4.4 Structures of the main core of VapG (molecule 1).

Secondary structure elements are labelled. The colour scheme indicates progression through the protein structure from the *N*-terminus (blue) to *C*-terminus (red). Figure was created using CCP4mg (McNicholas *et al*., 2011).

The core of the protein consists almost entirely of hydrophobic side chains from alternate residues of the β-strands which are densely packed, allowing no space for internal water molecules. There is a non-uniform distribution of polar and hydrophobic residues in the surface of VapG (Figures 4.5-4.6). At the top (referring to the orientation in figure 4.4), there is a preponderance of hydrophobic residues while at the bottom, the loops are longer and protrude from the barrel in the form of an α-helix. This region is comparatively richer in charged and polar side chains (Figure 4.6).

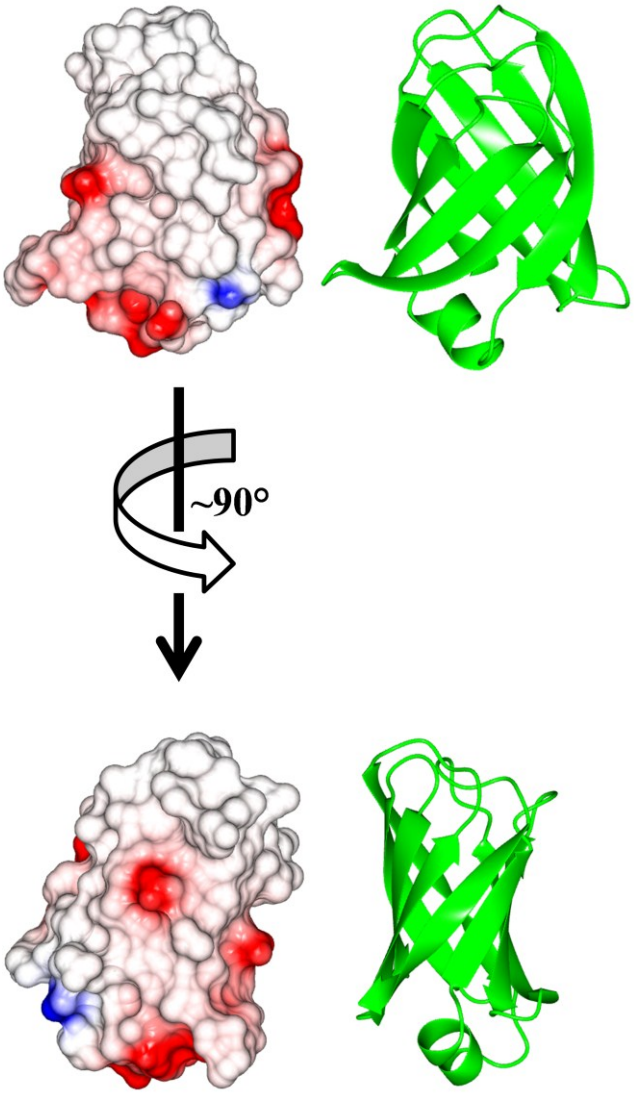


Figure 4.5 Electrostatic surface rendering of VapG (molecule 1) (A)
Positive and negative electrostatic potentials are shown in blue and red respectively. The apparent ribbon tracing are shown in green. The non-uniform distribution of the polar and apolar residues is evident. Figure was created using CCP4mg (McNicholas *et al*., 2011)

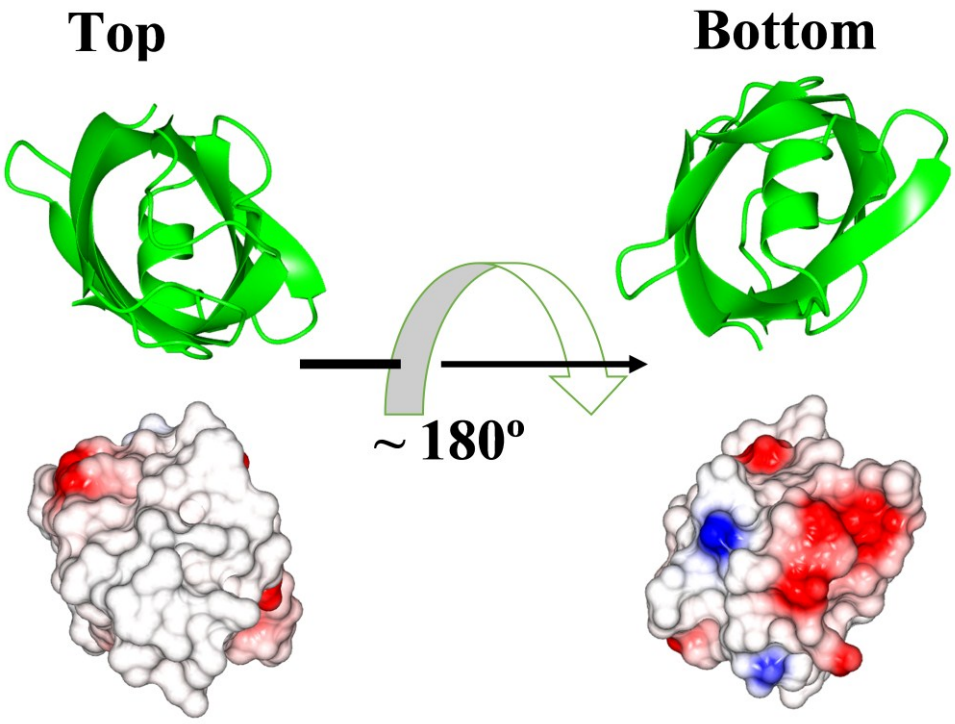


Figure 4.6 Electrostatic surface rendering of VapG (molecule 1) (B)
Positive and negative electrostatic potentials are shown in blue and red respectively. The apparent ribbon tracing are shown in green. Figure was created using CCP4mg (McNicholas *et al*., 2011)

The tight potassium coordination sphere of molecule 1 consists of six oxygen molecules. Two are contributed by the side chains of Asn107 and Asn114; three from the main-chain carbonyl oxygen atoms of Asn107, Val109, Tyr112 and the final ligand is a water molecule (Figure 4.7), which in turn H-bonds with the main chain amide of Tyr112.

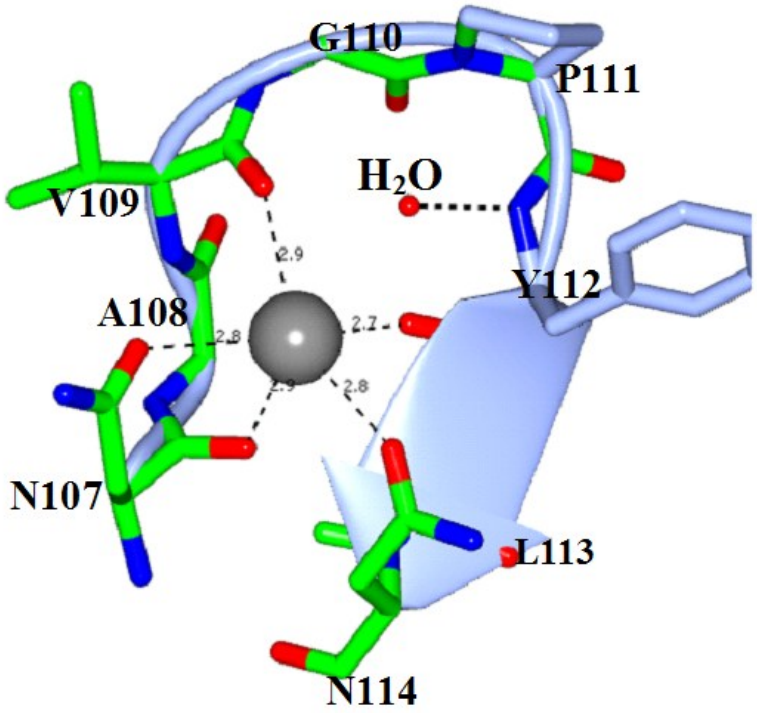


Figure 4.7 Metal coordination sphere of molecule 1.
Residues 107 – 114 which are involved in coordination with potassium (in grey) shown in cylinders. The water molecule (labelled) is represented by a red sphere. The residues are also shown in ribbon tracing (coloured in ice-blue) to reveal secondary elements. Figure was created using CCP4mg (McNicholas *et al*., 2011)

In summary, the VapG structure is similar to the VapD fold (Whittingham *et al.*, 2014) and the VapB fold (Geerds *et al*., 2014) (Figure 4.8). The only unique features of VapG are the presence of two molecules in the asymmetric unit and the possession of a coordinating sphere involving potassium. As this crystal was formed in a potassium-containing liquor, it is entirely possible that the metal binding is artefactual but was critical, nonetheless, in promoting crystallogenesis. It is interesting to note that poor crystals were also formed in liquors containing Li,

another group I metal but not with Na. These crystals did not produce high resolution data sets on diffraction so it is impossible to ascertain whether Li was bound in a similar fashion. Four of the coordinating groups where contributed by main-chain atoms or associated water, ultimately mutagenesis of Asn107 and Asn114 might inform our understanding of this metal binding and whether it has any significance to VapG function, presumably through seeking to measure a quantifiable contribution to virulence.

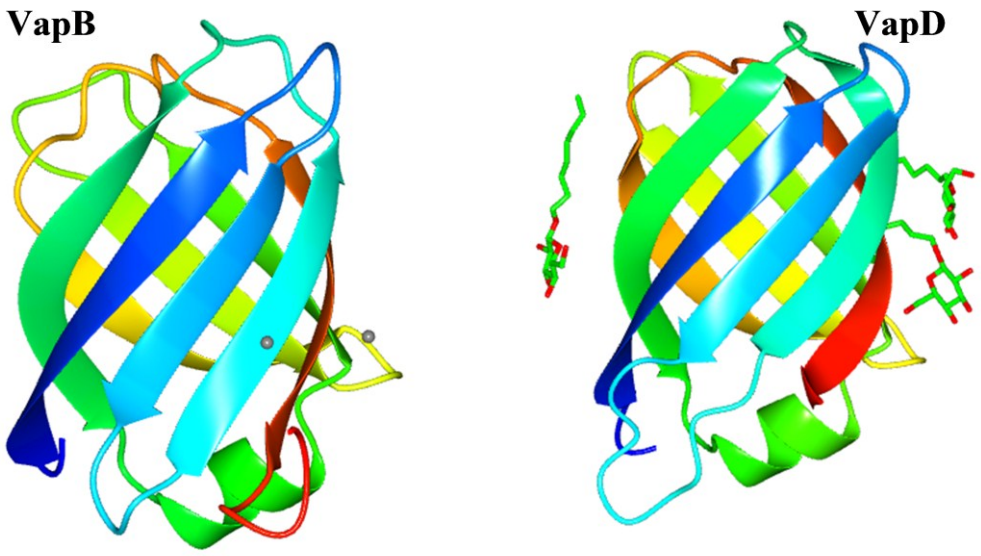


Figure 4.8 Structures of VapB and VapD colour ramped from the *N*-terminus (blue) to the *C*-terminus (red).
Structures were created using CCP4mg ((McNicholas *et al*., 2011) from the co-ordinates deposited in the Protein Data Bank as 4CV7 (VapB, Geerds *et al*., 2014) and 4CSB (VapD, Whittingham *et al*., 2014). Two Co^{2+} molecules are represented as spheres in VapB while three molecules of β-octylglycoside are represented as cylinders in VapD

The association of VapA with the mycolate layer that defines *R. equi* as a diderm is interesting. Several outer membrane proteins of Gram negative bacteria are also β-barrels (Wimley, 2003). In these organisms, this topology defines their localisation

and (probably) their function. These are often pore-forming proteins with consecutive β-strands forming a water-filled barrel in porins that facilitates transmembrane diffusion of various soluble materials (Pauptit *et al.*, 1991). Others have a plugged barrel in which a central globular domain occludes the central pore. This plug domain can afford a high-affinity binding site and direct the active transport route for key micronutrients like ferri-siderophores and cobalamin (Ferguson *et al*., 1998; Ratledge and Dover, 2000). Thus it might be tempting to speculate that VapG (and other vaps) may be directly associated with the mycolate layer of *Rhodococcus equi*. However, among the virulence associated proteins of *Rhodococcus equi* 103S, to date only VapA has been suggested to be tethered to the cell surface (Takai *et al*., 2000; Byrne *et al*., 2008).

VapG carries a few solvent-exposed polar residues at a predominantly apolar surface, a feature it shares with others defined (Geerds *et al*., 2014; Whittingham *et al.*, 2014) or modelled structures for *R. equi* Vaps (section 4.3). This polarity and its orientation are likely to have functional significance in the export or function of the secreted Vaps. The insight that these models have recently provided has delivered a blueprint for mutagenesis studies. This could eventually educate our understanding of Vap function when analysed in studies of *R. equi* virulence.

4.3 Homology modelling of other Vap Structures

4.3.1 Introduction

Despite the rapid increase in the number of known protein sequences, only about one-hundredth of these sequences have been characterised at high resolution using experimental structure determination methods (Webb and Sali, 2014). To bridge this

gap, comparative modelling remains the only method that can reliably predict the 3D structure of a protein with an accuracy that is comparable to that of a low-resolution experimentally determined structure (Moreira *et al*., 2013).

At the outset of this study, the *Rhodococcus equi* Vap family proteins were quite unique since there primary sequences were related to very few proteins in databases, presenting a challenge in building a structure using known models. However, recently the 3D structures of VapB and VapD have been resolved to a high resolution and deposited in the protein data bank with accession numbers 4CV7 and 4CSB respectively thus were considered as templates (depending on sequence similarity) for the modelling of the other Vaps. The proteins were modelled using Modeller, which is a python-based program developed by Sali (1995).

4.3.2 Structure modelling using Modeller

The steps involved in protein modelling using modeller have been described by Webb and Sali (2014). The process of model building consists of four main steps (Figure 4.9): (1) fold assignment that identifies overall similarity between the target sequence and at least one known template; (2) alignment of the target sequence and the template; (3) building a model based in the alignment of the chosen template and where applicable (4) predicting the accuracy of the model.

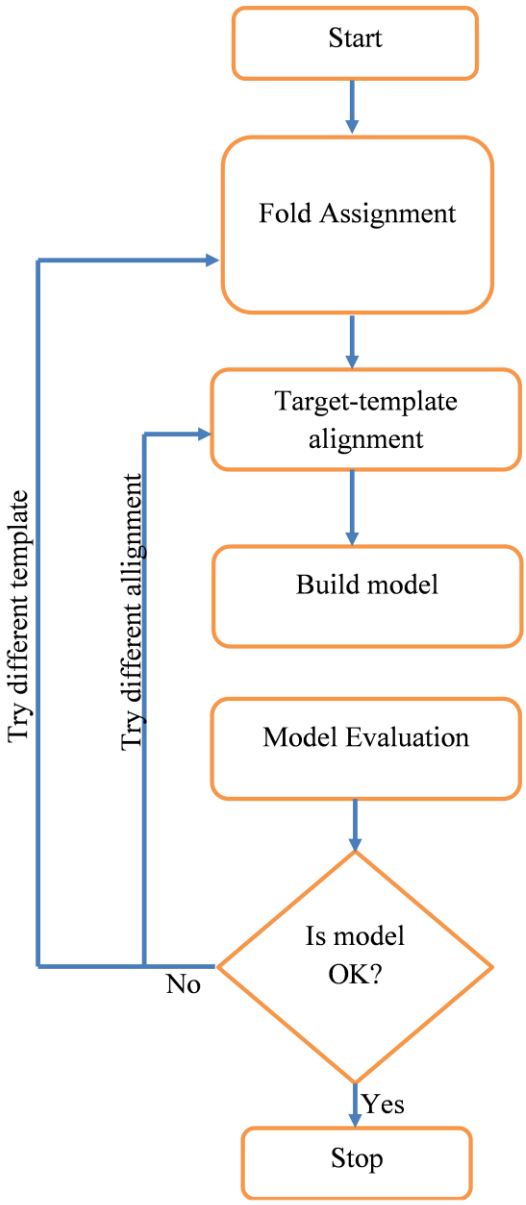


Figure 4.9: Flow chart illustrating the steps in protein modelling using modeller (adapted from Webb and Sali, 2014).

After installing the program, python scripts were written using the query sequence *i.e.* the amino acid sequence of the protein to be modelled, and the template *i.e.* the pdb file of the template available at the protein data bank. The scripts were run on

the program which led to the generation of the structure in pdb formats which are viewed after uploading on protein structure viewing programs.

4.3.3 Modelling of Vap proteins

Each of the amino acid sequence of the mature proteins (i.e. VapA, VapC, VapE, and VapH) were subjected to a blast search against the database of experimentally determined protein structures in the protein data bank to select the template for each protein. The search identified VapB and VapD with detectable similarity and the identity matrix is shown in table 4.2.

Table 4.2 Identity matrix between query sequences and template.
Statistics generated using BlastP.

		Potential Templates							
		VapB (Accession number 4CV7)				**VapD (Accession number 4CSB)**			
		Score	Query Cover	E value	Identity	Score	Query cover	E value	Identity
Models	VapA	169	68%	8e-54	87%	110	67%	1e-30	55%
	VapC	103	74%	3e-28	56%	92.4	73%	6e-24	50%
	VapE	96.7	71%	2e-25	52%	89.7	60%	1e-22	48%
	VapH	87.8	67%	4e-22	51%	77.4	66	4e-18	44%

In all cases, comparison with VapB produced more favourable scores for the queries than VapD, thus the 3D structure for VapB was used as a template for structural modelling of these proteins. The final modelled structures of the queries using VapB as a template are shown below in Figures 4.10-4.13.

The proteins contain a predicted cleavable signal sequence validated by SignalP (Appendix E) thus only the mature protein sequences were used in the modelling.

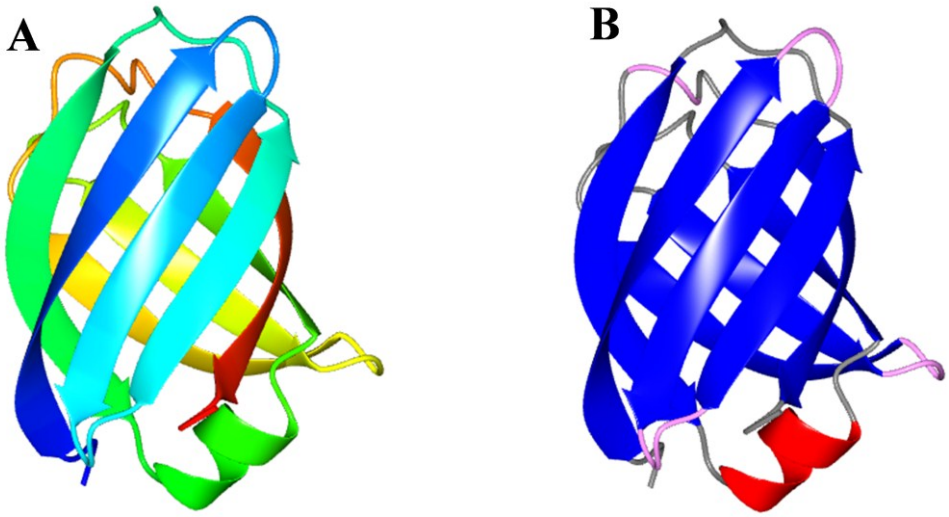


Figure 4.10 Model of VapA in cartoon mode
The model was created using CCP4mg (McNicholas *et al*., 2011) showing structure in A. Colour ramped from the *N*-terminus (blue) to the *C*-terminus (red). B. Coloured according to secondary structure elements (red, α-helices; blue, β-strands; others, loops).

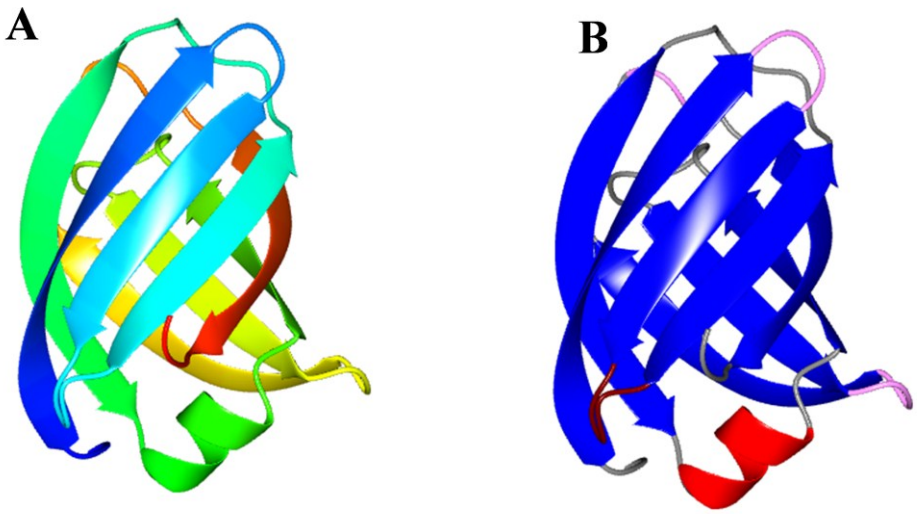


Figure 4.11 Model of VapC in cartoon mode
The model was created using CCP4mg (McNicholas *et al*., 2011) showing structure in A. Colour ramped from the *N*-terminus (blue) to the *C*-terminus (red). B. Coloured according to secondary structure elements (red, α-helices; blue, β-strands; others, loops).

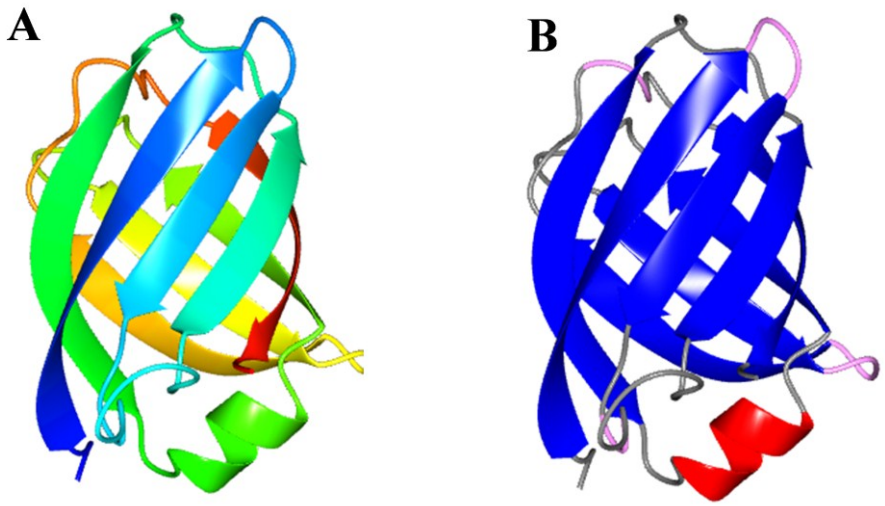


Figure 4.12 Model of VapE in cartoon mode
The model was created using CCP4mg (McNicholas *et al*., 2011) showing structure in A. Colour ramped from the *N*-terminus (blue) to the *C*-terminus (red). B. Coloured according to secondary structure elements (red, α-helices; blue, β-strands; others, loops).

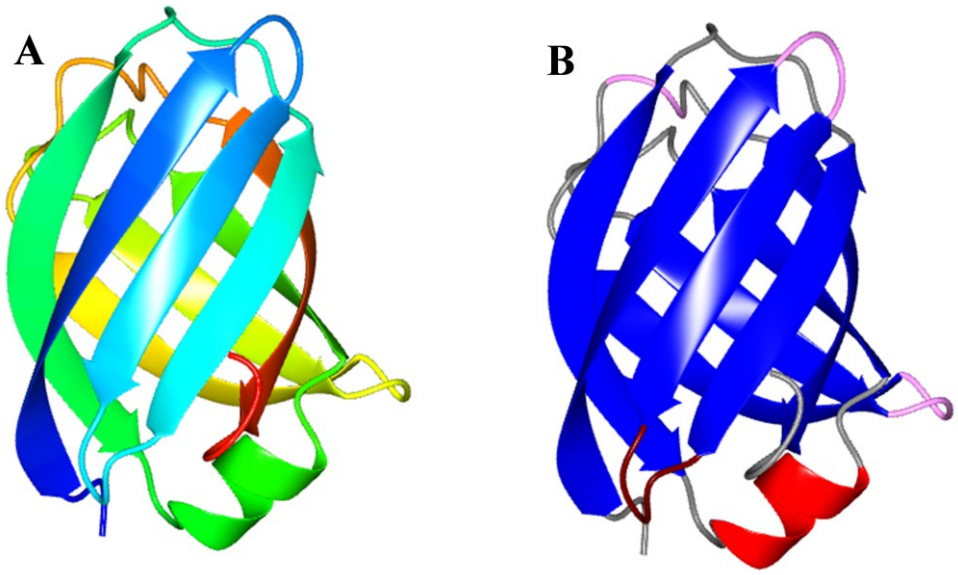


Figure 4.13 Model of VapH in cartoon mode
The model was created using CCP4mg (McNicholas *et al*., 2011) showing structure in A. Colour ramped from the *N*-terminus (blue) to the *C*-terminus (red). B. Coloured according to secondary structure elements (red, α-helices; blue, β-strands; others, loops).

The models conform well to the VapB fold (4CV7, Geerds *et al*., 2014), and are inherently similar to the structures of VapD (4CSB, Whittingham, 2014) and VapG (this chapter). Each model represents a compact 8-stranded β-barrel ordered as β1- β2- β3- β8- β5- β6- β7- β4 and an α-helix between β4 and β5 in the straight chain. The amino acids which make up each strand and helix are shown in table 4.3. These models superimposed well on the template structure, root means squared devations were low and ranged from 0.14Å to 0.38Å (Table 4.3).

Table 4.3 Structural elements of VAPs produced *via* modelling

Structural element	Template	Models			
	VapB	VapA	VapC	VapE	VapH
β1	Q88-A99	Q52-A63	E39-A50	K50-F61	Q48-A59
β2	V101-G109	V65-G73	F52-V60	F63-I71	V61-V69
β3	M113-G121	K77-G85	M64-G72	K81-G89	E73-G81
β4	G126-T136	G90-T100	G77-T87	G94-T104	G86-T96
α1	L139-D145	L103-D109	L90-E96	L107-E113	I99-R105
β5	V147-A153	V111-A117	A98-A104	V115-A121	E107-A113
β6	L158-D164	L122-D128	L109-D115	L126-A132	L118-D124
β7	G168-G176	S132-G140	A119-G127	G136-G144	A128-G136
β8	G183-H192	G147-H156	G134-H143	G151-R160	G143-I152
RMSΔ		0.14	0.38	0.24	0.25

Numbering of residues starts with 1 as the *N*-terminus of mature protein. The structure of the modelling template VapB is included for comparison. Average root mean squared deviations (RMSΔ) were calculated using PDBeFold (http://www.ebi.ac.uk/msd-srv/ssm/)

4.3.4 Model Evaluation

Since the models were produced theoretically, the quality of each model was evaluated based on stereochemical and geometrical restraints using COOT (Emsley *et al*., 2010) and ProSA-Web (Wiederstein and Sippl, 2007). COOT was used for the generation of Ramachandran plots (to assess backbone dihedral angle constraints) (Figure 4.14) while ProSA-Web was used for the generation of the quality score of

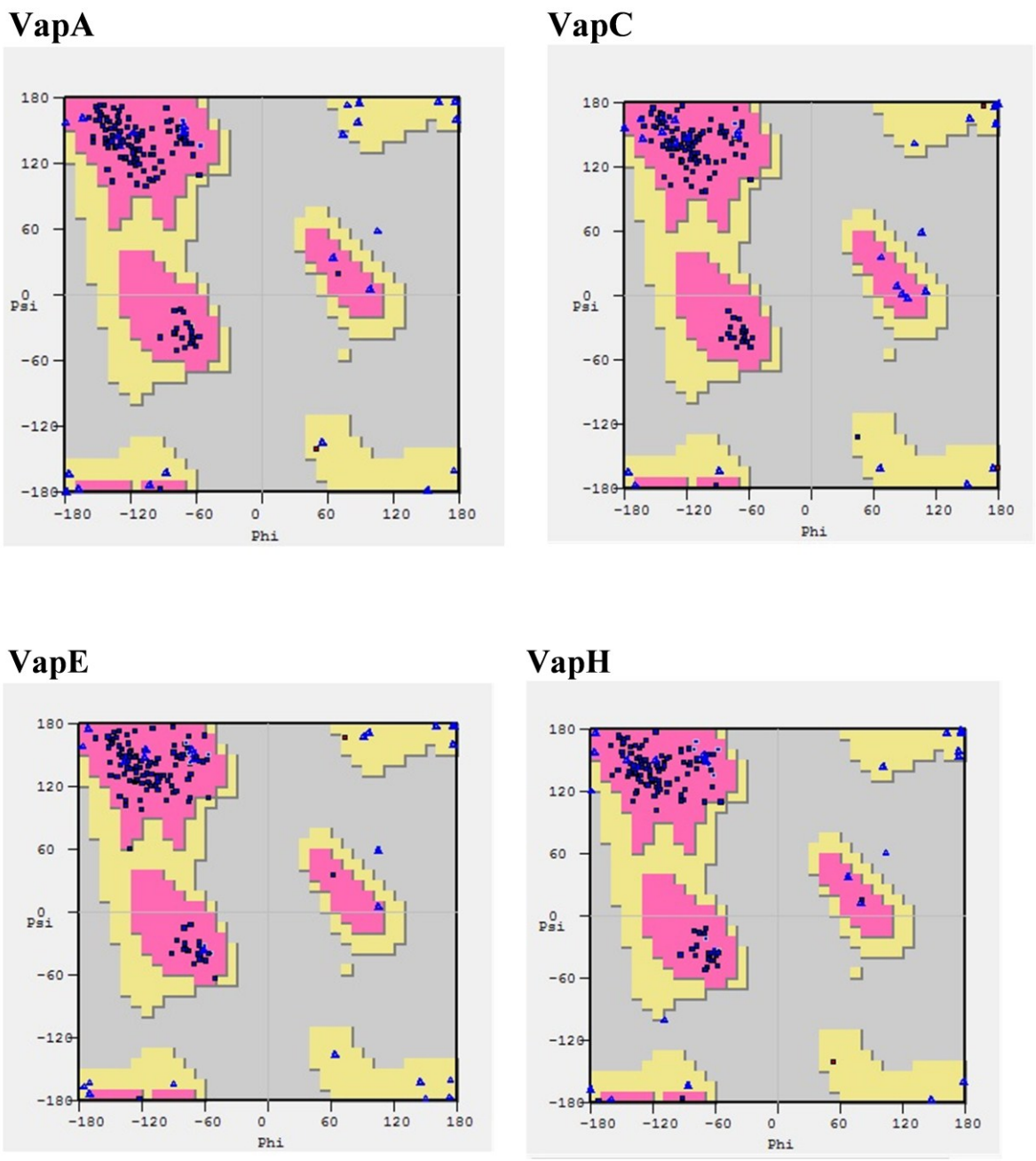


Figure 4.14 Ramachandran plots for the models for VapA VapC, VapE and VapH.

Plots were generated by COOT (Emsley *et al*., 2010). The regions are coloured coded to indicate combinations of dihedral angles considered as preferred (pink), allowed (yellow) and outliers (grey).

Table 4.4 Summary of Ramachandran statistics and ProSA Z-scores for the Vap models.

Model	Prefered Region (%)	Allowed Region (%)	Outliers	Pro Z-score*
VapA	97.44	1.92	0.64	-4.4
VapC	96.50	2.10	1.40	-4.35
VapE	98.10	1.27	0.63	-4.88
VapH	95.45	3.90	0.65	-3.35

Plots for the Z-scores (including VapB) using ProSA-web are in appendix G.

each model against the scores of protein structures obtained from X-ray analysis and NMR spectroscopy (Table 4.4).

As shown in table 4.4, the Ramachandran statistics reveal that the residues which have psi/phi angles in the most favoured regions range from 95.45 % (in VapH) to 98.1 % (in VapE) while 1.27 % (in VapE) to 3.90 % (in VapH) of the residues fall in the allowed regions. The residues that fall in the disallowed regions range from 0.63 % (in VapE) to 1.40 % (in VapC). The Z-score when the pdb file for VapB was analysed *via* ProSA-web was -4.34 while the Z-score for the models range from -4.88 to -3.35. This shows that the Z-score for the modelled proteins are very similar to the experimentally-determined structures of proteins of similar size and this suggests further reliability to these 3D-models (Appendix G).

4.3.5 Model orientation and surfaces

The apparent strong conservation of the fold within this protein family indicated by this modelling study suggests that and either (i) these proteins are very similar and there is great functional redundancy encoded within each PAI, (ii) any functional resolution is defined in the unstructured *N*-terminal regions, (iii) surface chemistry in the structured core dictates function or (iv) combinations of (ii) and (iii) are

responsible for individual contributions to virulence. Using these models, (iii) could be considered. Figure 4.15 illustrates the electrostatic surface rendering for each of the four models. It is clear that the general trends observed earlier for the crystal structure of VapG are evident here. The 'top' of the molecules have a more hydrophobic nature; polar residues are distributed across the bottoms and the lower half of the barrel walls. Beyond these general principles however, it is clear that the patterns of charge distribution exhibited by the individual proteins are distinctive and could represent the structural basis for a range of specific interactions that promote virulence. These molecular interactions might include those that could be categorised *via* the origin of the partners as Vap:Vap, Vap:other bacterial or Vap:host. These models, along with the three experimentally-derived Vap structures, provide a blueprint for site-directed mutagenesis studies to dissect the structure: function relationships of these interesting β-barrel domains.

4.4 The unstructured N-terminus

One crucial feature of each of the 3D crystal structures defined to date is the absence of structure in the *N*-terminal regions of VapB, VapD and VapG. This region represents about 30 % of each molecule and could be important functionally to this class of proteins since it is structurally conserved. Although these regions (intrinsically disordered or unstructured regions as they are sometimes called) do not form well-defined structures in their native states under crystallisation conditions, they are likely free to rapidly sample many alternate configurations and could form stable 3D- structures upon binding to specific molecular targets (Huang and Sarai, 2012). This process, which is referred to as coupled folding and binding, is facilitated due to their flexibility and is involved in a number of processes such as

phosphorylation, transcription control and also disease development (Fukuchi *et al*., 2011, Ota *et al*., 2013). Bulky and aromatic amino acid residues are order-promoting due to their ability to form the hydrophobic core of folded globular proteins (Uversky, 2013). These order promoting groups appear to be limiting in the long unstructured regions of the Vaps. An inspection of the VapA structure (model) reveals that over 94 % of aromatic amino acids residues (e.g. trytophan, phenylalanine and tyrosine) are found in the β-barrel core. However, there are twelve serines (which are predominantly clustered) in the long unstructured part of the model. Serine is a disorder-promoting residue (Uversky, 2013). Thus the low mean hydrophobicity favours the formation of the unstructured region of VapA (Uversky, 2011). Though this feature appears to be least conserved among VapB, VapD and VapG, it could be an important element *via* which interaction with host cellular targets, other Vaps or other bacterial components are mediated to promote the virulence of *Rhodococcus equi*.

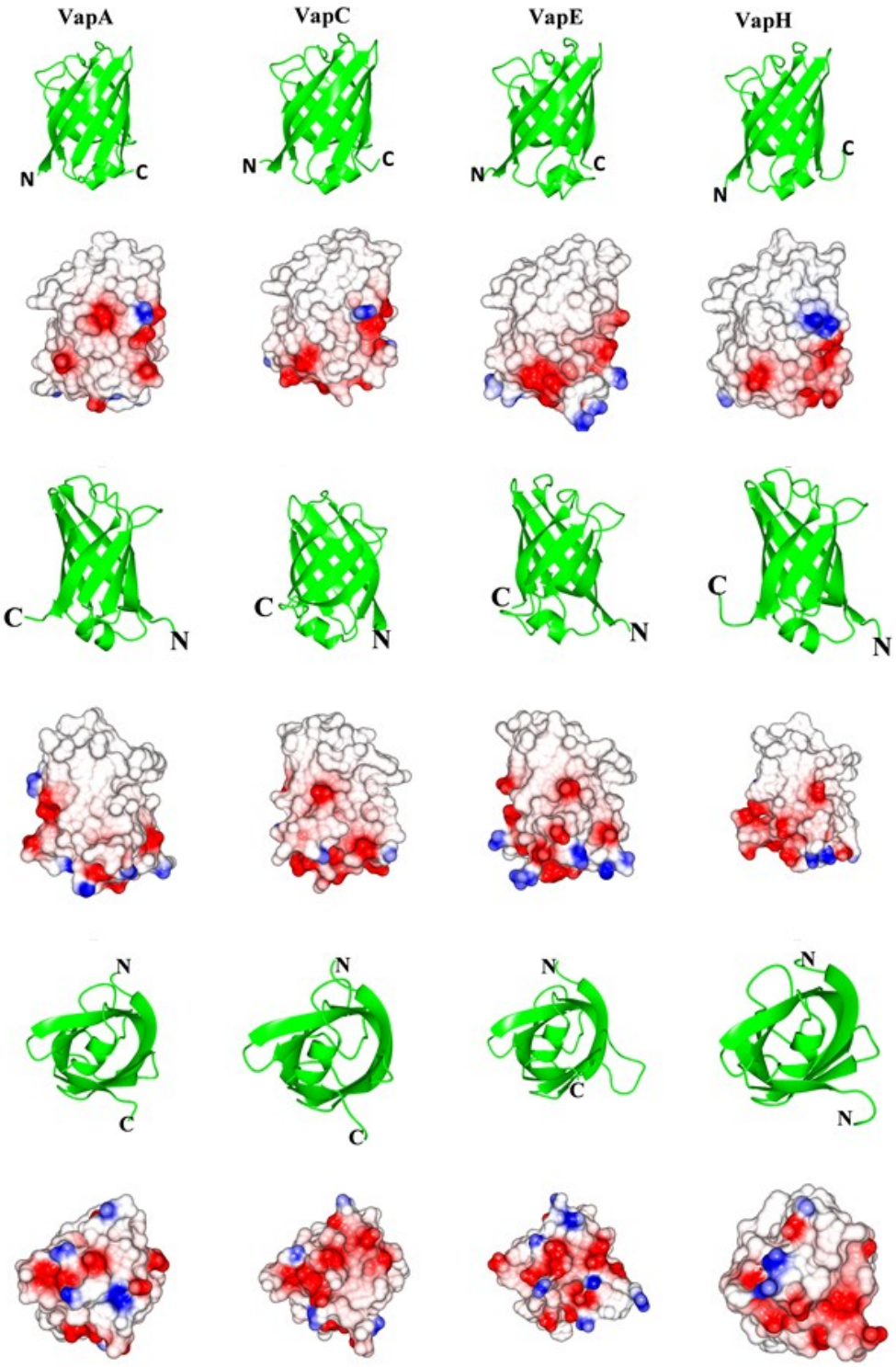


Figure 4.15 Electrostatic surface rendering of Vap models
Positive and negative electrostatic potentials are shown in blue and red respectively. The apparent ribbon tracing are shown in green.

CHAPTER FIVE

5 Interactions of virulence associated proteins

5.1 Introduction

Almost all biological processes are either direct or indirectly controlled by macromolecules, especially proteins. Association between proteins plays a central role in the structural assembly and functional regulation of biological systems (Freire *et al*, 1990; Ladbury and Chowdhry, 1996). Processes such as nastic responses, gene expression, polarisation/depolarisation of action potentials, hormone-receptor interactions and cell-ligand interactions, amongst others, are mediated by specific protein-protein interactions. These interactions involve dipole-dipole effects, covalent bond formation, hydrophobic interactions and the formation of salt bridges and are characterised by some thermodynamic variables like enthalpy, entropy, free energy and stoichiometry of the interacting species.

Even though the precise role of the Vap proteins is yet to be elucidated, VapA has been reported to be a possible virulence determinant from gene knock-out studies. Jain *et al*. (2003) reported that an *R. equi* mutant lacking a 7.9 kb DNA region spanning five *vap* genes (*vapA*, -*C*, -*D*, -*E* and –*F*) was attenuated for virulence in mice as it was unable to replicate *in vivo* and was rapidly cleared by the immune system. In complementation analyses, expression of *vapA* alone restored full virulence, whereas expression of *vapC*, -*D* and -*E* did not. Similarly, an *R. equi* strain lacking only *vapA* was also attenuated for growth *in vivo*. Both of these mutant strains could attach to but were unable to grow in macrophages. Although these studies established the importance of VapA, others have demonstrated that the protein is not sufficient for the proliferation of *R. equi* in macrophages (Giguere *et al*,

1999; Coulson *et al*, 2010). A recombinant, virulence plasmid-cured strain that expressed *vapA* to wild-type levels did not grow in macrophages and was avirulent in experimental infections of both mice and foals (Giguere *et al*, 1999). Similarly, Coulson *et al*. (2010) constructed an avirulent PAI deletion mutant. When *vapA* was expressed from a constitutive promoter in this mutant, full virulence was not restored.

It is clear that the PAI of the *R. equi* virulence plasmid is essential for intra-macrophagic replication and that VapA plays a major role in this, but VapA alone is not sufficient to restore virulence indicating that additional virulence genes are located within the pathogenicity island.

The production of VapA is coordinated with other Vaps. Both VapA and VapD are induced in response to acid stress while the transcription of both *vapA* and *vapG* increases in the presence of hydrogen peroxide. Temperature and pH could also up-regulate gene expression in a synergistic manner (Miranda-CasoLuengo *et al*, 2011). This data may suggest that VapA, VapD and VapG may act in concert to enable the survival of *Rhodococcus equi* in macrophages (Benoit *et al*, 2002). Though the interactions between the Vap proteins are not known, available evidence seem to suggest that they all may contribute to the virulence of the organism based on their homology, regulatory strategies and the presence of their genes on a single locus (Jacks *et al*, 2007; Letek *et al*, 2008).

We hypothesised that the importance of VapA may reside in its ability to be retained at the cell surface by its unusual lipid modification and that other Vaps might interact with it to promote virulence as an assemblage. Potential interactions of Vap proteins were explored and investigated through the use of experimental biochemical and

biophysical approaches. Specifically, a combination of electrophoresis in native gels, analytical ultracentrifugation (AUC) and surface plasmon resonance (SPR) were deployed to follow Vap interactions *in vitro*. All of the data described in this chapter is derived from the use of *C*-terminal tagged Vaps.

5.2 Self-association and the influence of metals

VapA is recognised for its complex electrophoretic migration. The characteristic broadening of the VapA band might be related to its unusual lipidation or an ability to interact with other migrating species or with the polyacrylamide matrix itself even under the denaturing conditions of SDS-PAGE. This broad banding pattern was diminished but not completely abrogated in this study when an *N*-terminally truncated (lacking its signal sequence) *vapA* was expressed in a heterologous host. An unusual lipidation is less likely in this situation but the possibility of retarding interactions remains. One distinct possibility is that VapA might oligomerise. We considered the possibility that any potential self-interaction might be enhanced under native electrophoresis conditions. When VapA was tested, two distinct bands were observed that could most easily be reconciled as monomeric and dimeric forms of the protein (Figure 5.1). Evidence for self-association of VapG was not observed using this methodology.

Figure 5.1 Purified VapG and VapA subjected to Native-PAGE.

When early data were emerging from the analysis of VapG diffraction data, the possibility that this protein might bind a metal, likely Ca^{2+} or K^{+} was envisaged; a property not associated with VapD (Pers. comm. A. Wilkinson); no data were available to us for other Vaps in this regard. In order to investigate whether this metal binding might impact upon the ability of VapG to associate with itself or other Vaps, the protein was incubated in the presence of each metal (1 – 10 mM) for 30 mins and migration on native gels was analysed, no evidence was found for alteration of its electrophoretic migration in isolation (data not shown). Similar experiments with VapA revealed that its apparent dimerisation was not affected in the presence of these metals (data not shown).

With reference to our working hypothesis that other Vaps might associate with VapA, mixtures of VapA and VapG were analysed in the presence and absence of these two metals. Again no evidence was seen for an association of these two Vaps in either scenario (data not shown). However, this methodology is crude and the

interactions explored were limited; any physiological association of VapA and VapG might be indirect and mediated by another agent. However, clear evidence for VapA dimerisation was observed so confirmation was sought *via* an alternative methodology.

5.3 Analytical ultracentrifugation

Analytical ultracentrifugation is an important technique used to characterise macromolecular associations (in terms of stoichiometry and strength) in solutions (Howlett *et al*, 2006). The following section describes the use of AUC to study the sedimentation properties of VapA and VapG and the effect of metals on self-association of these Vaps. Experiments were carried out in The University of Birmingham Institute of Microbiology and Infection as a guest of Prof. Gurdyal Besra.

In order to validate the possibility of self-association, each purified recombinant protein (VapA and VapG) were subjected to analytical ultracentrifugation at 40,000 rpm in the presence and absence of metals (KCl, LiCl or $CaCl_2$ at 1mM) using the sedimentation velocity mode. LiCl was included in the study at this stage as poor quality VapG crystals had formed in Li^+-containing liquors. Since it was considered that the binding of a Group I metal might be a pre-requisite for crystallogenesis, crystal formation in Li^+ might indicate a valid interaction. The absorbance data for the effect of the metals on the sedimentation behaviour of VapA is shown in figures 5.2, while the data on the effect of the metals on VapG is shown in figures 5.3 and 5.4 and summarised in Table 5.1.

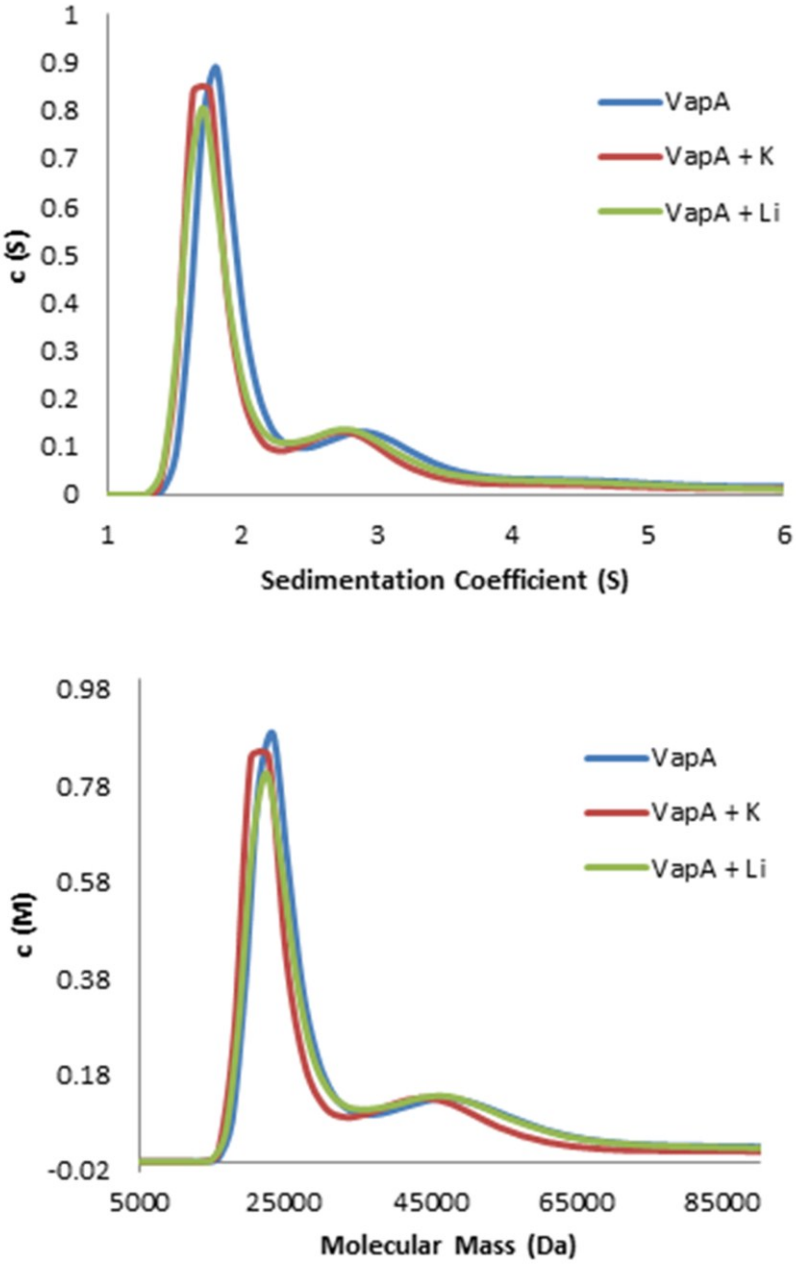


Figure 5.2 Influence of metals on sedimentation of VapA in analytical ultracentrifugation

Recombinant VapA was incubated with 1mM KCl or 1mM LiCl for several minutes prior to centrifugation. Both VapA and salts were dissolved in 0.02 M Tris.HCl pH7.9, 0.05 M NaCl. The upper panel describes the relationship between the absorbance signal (c(s)) and sedimentation coefficient (S) and the lower panel relates absorbance signal (c(M)) to the apparent molecular mass.

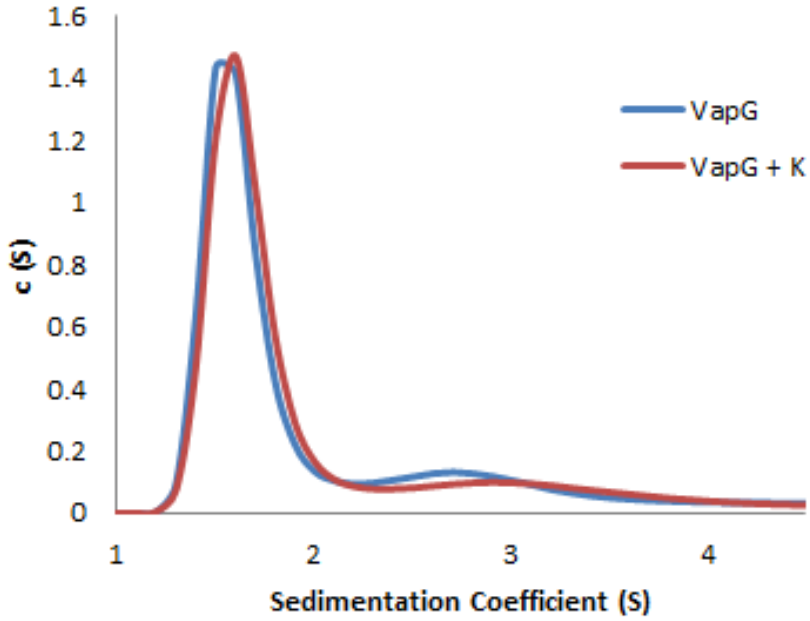


Figure 5.3 Influence of 1mM potassium on the sedimentation coefficient of VapG in analytical ultracentrifugation
Both protein and salt were re-suspended in 0.02 M Tris.HCl (pH7.9 containing 0.05 M NaCl)

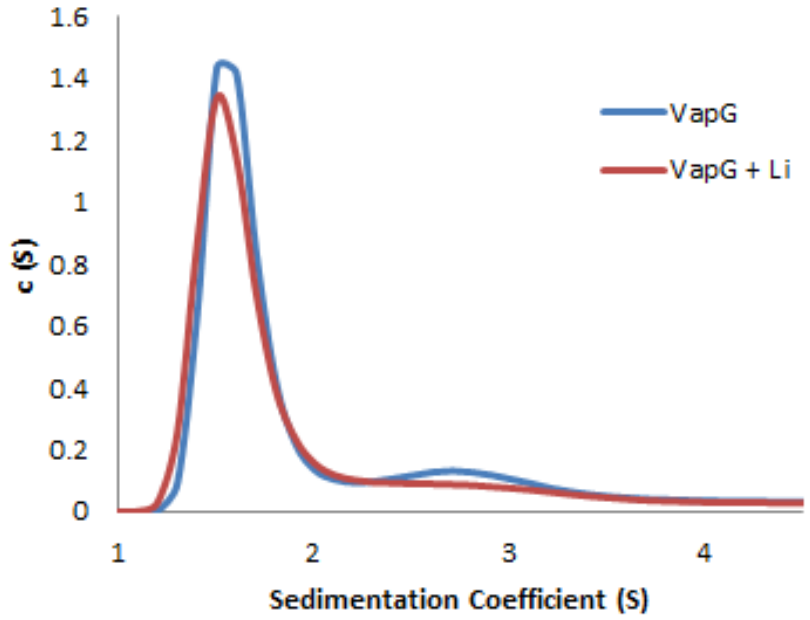


Figure 5.4 Influence of 1mM lithium on the sedimentation coefficient of VapG in analytical ultracentrifugation
Both protein and salt were re-suspended in 0.02 M Tris.HCl (pH7.9 containing 0.05 M NaCl)

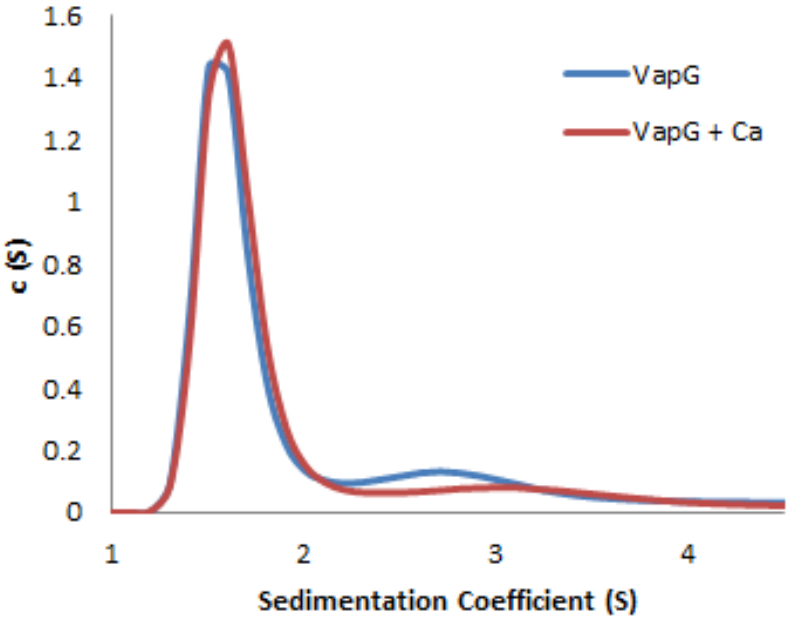


Figure 5.5 Influence of 1mM calcium on the sedimentation coefficient of VapG in analytical ultracentrifugation
Both protein and salt were re-suspended in 0.02 M Tris.HCl (pH7.9 containing 0.05 M NaCl)

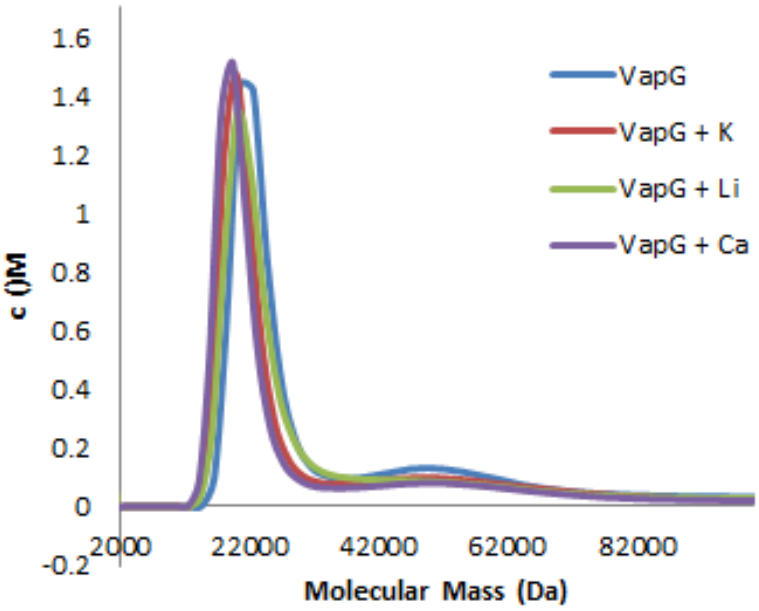


Figure 5.6 Influence of 1mM metals on the apparent molecular mass of VapG in analytical ultracentrifugation
Both protein and salts were re-suspended in 0.02 M Tris.HCl (pH7.9 containing 0.05 M NaCl)

This data reveal that both VapA and VapG were predominantly monomeric but did not sediment in a manner expected of a globular protein of the calculated molecular mass of the recombinant molecule. These masses were over-estimated in both cases by approximately 6 and 4.6 kDa for VapA and VapG respectively. Long unstructured (disordered) motifs of proteins have larger hydrodynamic dimensions than globular equivalents of the same molecular mass (Kozlowska *et al*, 2014). VapG has a β-barrel structure and possesses a long unstructured region at the *N*-terminus (Chapter 4). VapA shares significant homology with the structurally characterised VapB, which also shares this characteristic (Geerds *et al*., 2014). Thus these unstructured regions of the proteins could be a major contributing factor to the difference between the apparent and the calculated molecular masses of the recombinant molecule on AUC.

However with the increasing protein concentration associated with sedimentation, a second molecular weight became apparent which indicates the formation of a dimer. The observations reveal that both VapA and VapG could exist in a monomer-dimer equilibrium at high concentrations. These equilibria would require multiple AUC runs in sedimentation equilibrium mode to resolve. A superficial examination of these traces does reveal that the ratio of peak heights (monomer: dimer) for proteins in the absence of metal were 6.8 and 10.9 for VapA and VapG respectively. This is not a rigorous examination but it suggests that VapA has a greater predisposition for dimerisation than VapG, which is in agreement with the dimerisation seen on Native-PAGE (Figure 5.1) for VapA only.

The presence of metals in the medium did not change the sedimentation behaviour of VapA significantly as both signals for monomer and dimer were retained (Figure

5.2). However, the dimer signal (centred around 2.68 S, Figures 5.3-5.5) for VapG seem to have been lost with the metals in the medium. It appears that exposing VapG to the metals reduced the ability of protein to self associate (*i.e*. homodimerise) even with increasing protein concentration upon sedimentation.

Table 5.1 Sedimentation properties of VapA and VapG

		VapA (17.6 kDa)		VapG (16.8kDa)	
	Metals	*Monomer	*Dimer	1	2
Sedimentation Coefficient (S)	Nil	1.8	2.9	1.6	2.7
	K^+	1.7	2.8	1.6	0
	Li^+	1.7	2.8	1.5	0
	Ca^{2+}	N/A	N/A	1.6	0
Molecular Mass (kDa)	Nil	23.6	46.9	21.4	48.5
	K^+	21.6	44.7	19.7	0
	Li^+	22.3	45.9	20.2	0
	Ca^{2+}	N/A	N/A	19.5	0

All values are approximations based on data analysis fitting a single frictional coefficient. N/A, effect of Ca^{2+} on sedimentation behaviour of VapA was not performed. *These are apparent values from sedimentation velocity experiments. Values in brackets are the calculated molecular masses of the recombinant molecule using ExPASy Protparam.

5.4 Surface plasmon resonance

Although AUC could be used to carry out studies of Vap interactions, this was not convenient and uses high concentrations of protein samples. We chose to explore potential interactions between Vaps using surface plasmon resonance (SPR), a flexible label-free technique through which both affinity and binding constants can be obtained from binding responses over a range of concentrations (Besenicar *et al*, 2006) in real time using relatively small amounts of analytes in a native or native-like environment (Stahelin, 2013; Patching, 2014). This latter aspect was interesting as interactions between immobilised VapA and other Vaps could be measured at an interface, to an extent emulating the *in vivo* scenario. SPR was available locally and

experiments were carried out as a guest of Prof Jeremy Lakey at the University of Newcastle upon Tyne.

5.4.1 Interaction of VapA with other Vaps

In *Rhodococcus equi*, available evidence suggests VapA has an unusual lipidation and is tethered to the cell wall surface while other Vaps are mostly secreted and diffusible proteins (Byrne *et al*., 2008). It is possible that other Vaps might associate with VapA to form an assemblage that might promote virulence. Based on this intrinsic property of VapA to adhere to the cell surface, it was chosen to be immobilised on the SPR sensor chip as the ligand; the experimental model involved the passage of other Vaps over the immobilised VapA to monitor interactions that might promote such assembly.

VapA was adsorbed on a Biacore CM5 sensor chip (*via* amide coupling as described in section 2.8.16). After 1 min of loading VapA (at 42 µg/mL) at 5 µl/min, a signal of 1483 RU was obtained which was stable after completion of the blocking protocol to prevent covalent capture of analytes by the activated carboxymethylated dextran matrix.

For analytical runs, solutions of the other Vaps available to us VapD, E, G and H were passed over the VapA chip for 3 minutes in an association phase at one of two concentrations (50 µM and 100 µM in HBS buffer; 0.01 M HEPES, 0.15 M NaCl, 0.003 M EDTA, 0.05 % P20). The analysis was continued in a dissociation phase in which HBS buffer was passed over the chip. Data were adjusted by subtraction of the signal from the analyte passing simultaneously through the control flow-cell. Adjusted and compiled sensograms are shown in figure 5.7.

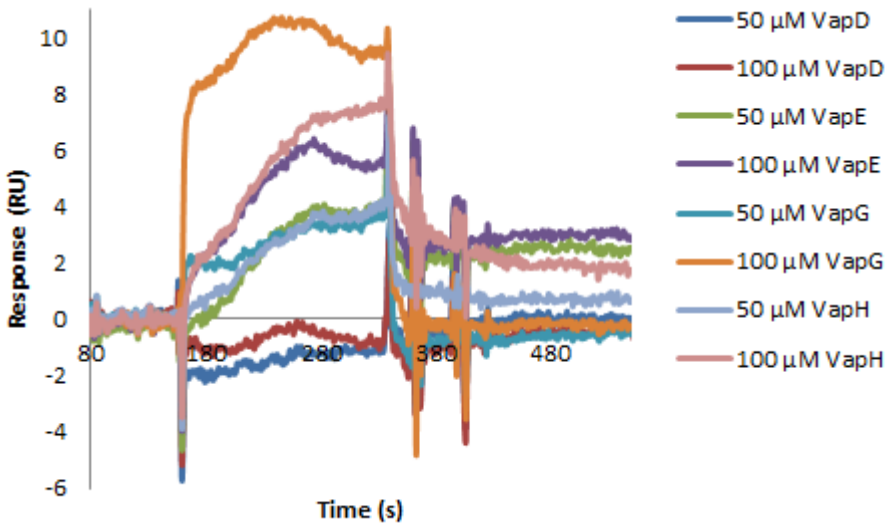


Figure 5.7 Interaction of *Rhodococcus equi* virulence associated proteins with immobilised VapA
Individual Vaps were run as analytes on a VapA modified CM5 chip at 50 µM and 100 µM using a Biacore X100. 1434 RU VapA was bound to the CM5 sensor chip.

The experiments with VapD resulted in no measurable response indicating a very poor binding between VapA and VapD in this format. Experiments with VapE and VapH produced very similar responses, showing a gradual accumulation of signal over approximately 2 minutes, which continued for at a slower rate for the rest of the association phase in most cases, VapE at 100µM showed a minor loss of signal at this point. The magnitude of the apparent binding was broadly concentration dependent. VapG behaved differently with complex concentration dependence and a markedly more rapid association in the early part of the experiment. On entering the dissociation phase of the experiment, the signal deteriorated in the order VapG>VapH>VapE.

As VapG produced the strongest signal at the concentrations considered, a broader range of VapG concentrations between 1 and 100µM was considered (figure 5.8).

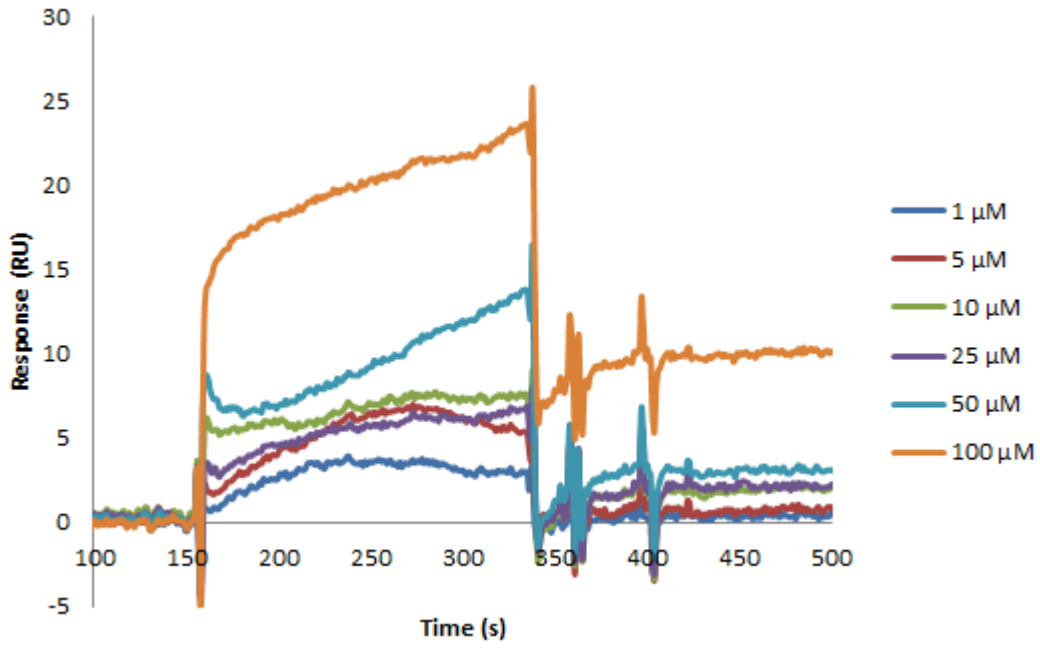


Figure 5.8 Interaction of *Rhodococcus equi* VapG with immobilised VapA
VapG was run as an analyte on a VapA-modified CM5 chip at concentrations between 1µM and 100 µM using a Biacore X100. 1434 RU VapA was bound to the CM5 sensor chip.

Significant binding of VapG to VapA was observed at all concentrations thus the kinetics of binding (including binding of VapE to VapA) was measured using the 'wizard' mode, which manages automatic runs with in-built technical replication.

5.4.2 Binding kinetics

The kinetics of binding of both VapG and VapE to immobilised VapA were considered using various concentrations between 1µM and 100 µM *via* the automated 'wizard' mode. Data were fitted using the 1:1 interaction model which assumes a simple reversible biomolecular interaction between the ligand *i.e.* VapA

and the analyte *i.e.* either VapG or VapE which can be represented by the following equation.

$$\text{VapA}_{(\text{immobilised})} + \text{Ligand (VapG or VapE)} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \text{VapA:ligand}$$

Where the association rate constant and dissociation rate constant are represented by k_a and k_d respectively). Due to mass transport effects that were envisaged, ligand concentrations higher than 100 µM were not used since it could be rate limiting thereby reduce the reliability of the binding data (Rich and Myszka, 2000; Tudos and Schasfoort, 2008).

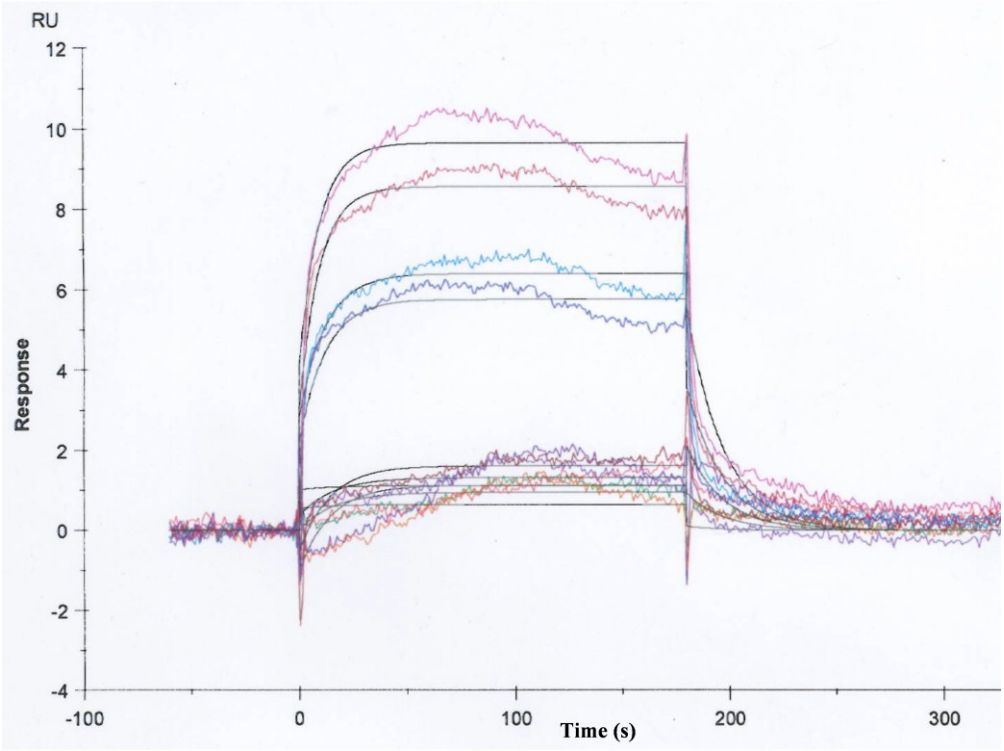


Figure 5.9 Kinetic analysis of VapG with immobilised using surface plasmon resonance.
Each concentration (1 µM, 10 µM, 25 µM, 50 µM and 100 µM) was run in duplicate using a Biacore X100. Prior to fitting, values for the blank *i.e.* HBS running buffer were subtracted from real experimental values. Fitted curves are superimposed in black.

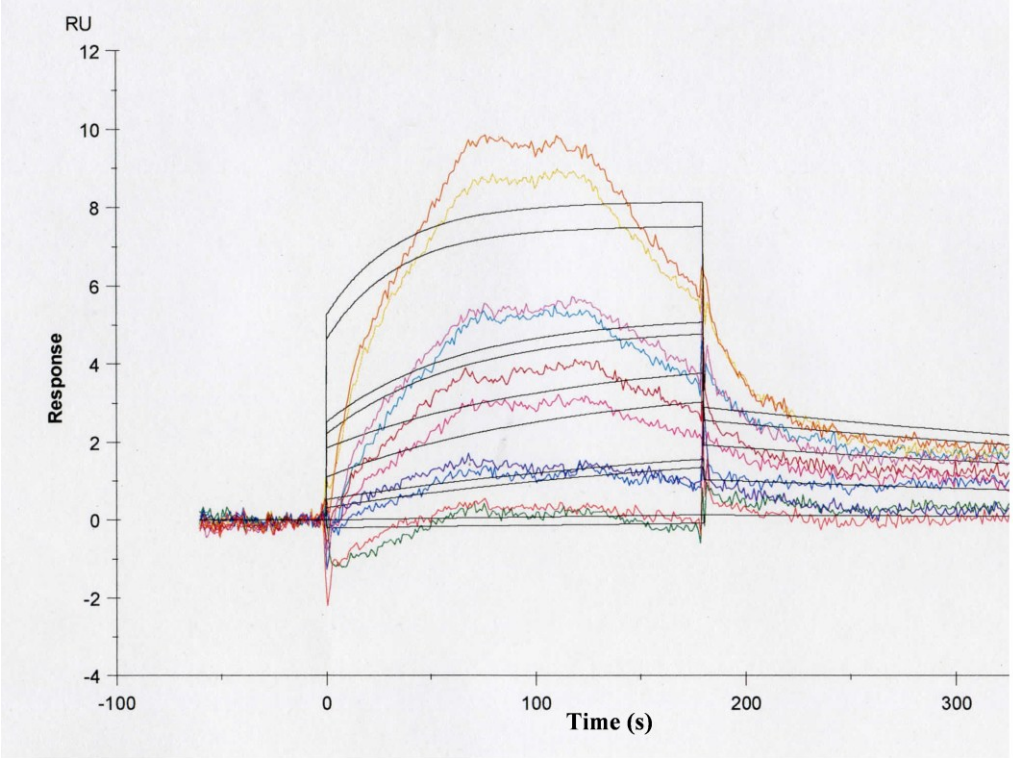


Figure 5.10 Kinetic analysis of VapE with immobilised VapA using surface plasmon resonance

Each concentration (1 µM, 10 µM, 25 µM, 50 µM and 100 µM) was run in duplicate using a Biacore X100. Prior to fitting, values for the blank *i.e.* HBS running buffer were subtracted from real experimental values. Fitted curves are superimposed in black.

After measuring the interactions *via* the wizard mode (*i.e.* VapG/E interaction with VapA), the data was fitted using a 1:1 interaction model using the Biacore X100 evaluation software. In the case of VapG, the experimental data fitted the 1:1 binding model very well (Figure 5.9) with the calculated descriptive constants presented in Table 5.2. The fit for VapE was less convincing. However, a generally accepted indicator of quality of fit for SPR data is indicated by the ratio of R_{max} and Chi^2 generated by the Biacore X100 Evaluation Software. Generally, a Chi^2 amounting to

less than 10% of R_{max} is accepted for a fit (Yapeng *et al*., 2013). In this respect both data sets would be acceptably described by this 1:1 binding model.

Table 5.2 Kinetic parameters for VapG/E interaction to VapA from a 1:1 binding model

Parameter	VapG	VapE
K_a (1/Ms)	471.7	272.0
K_d (1/s)	0.05895	0.001951
K_D (M)	1.25×10^{-4}	7.17×10^{-6}
R_{max} (RU)	12.64	3.119
Chi^2 (RU^2)	0.129	0.264

VapG was found to interact with VapA with an equilibrium dissociation constant (K_D) of 125 μM while the K_D for the interaction of VapE to VapA was 7.17 μM (Table 4.3). This reveals a weak interaction between VapG and VapA *i.e.* K_D > 100 μM and a higher affinity interaction (Qin and Gronenborn, 2014) between VapE and VapA. Thus VapE is more readily recognised by VapA thus a VapAE complex is more stable than a VapAG complex.

VapA was immobilised on the sensor chip *via* lysine residues and there are three lysines in the mature protein sequence. In the 3D-model of VapA, Lys 37 is part of the unstructured *N*-terminus, Lys 77 is in the β3 while Lys 108 is part of the α-helix. The side chain of Lys 108 is outside the core so Lys 37 and108 could be solvent-exposed and available for reaction with the activated carboxymethylated dextran matrix. The side chain of Lys 77 is buried inside the core so may not be solvent exposed. Thus two modes of immobilisation could be reflected in these experiments. Steric effects could limit interactions in either case. It is likely that the unstructured *N*-terminus is involved in binding events; the immobilisation of VapA *via* Lys 37 might severely restrict the mobility of this region and limit its participation in any

possible Vap-Vap interactions. Likewise immobilisation *via* Lys 108 might also restrict access of binding partners to important sites. Another mode of immobilisation could be attempted using the hexahistidine tags (either *N*-terminal or *C*-terminal) of the recombinant VapA proteins on Ni-NTA chips. In each case the orientation of the protein would be better understood and differential affinities that might arise for each Vap analyte might be rationalised. Here analyte proteins would need to be untagged. Additionally, the individual contributions of the VapA core and its unstructured *N*-terminal region could be analysed separately by producing discrete truncated molecular constructs to be immobilised on Biacore chips.

Another factor that should be explored is the ambient pH at which such interactions might occur. One factor that is particularly relevant is the up-regulation of *vapA* and *vapD* at acidic pH (Benoit *et al*., 2001); superior results might be achieved by conducting these binding experiments at a range of pH values.

This preliminary study of Vap interactions is limited in its scope but has revealed some interesting avenues for further study. It is clear that Vaps might associate with VapA to varying extents and with varying affinities. The value of this method should be evaluated further.

CHAPTER SIX

6 Lipidation of Envelope proteins and recombinant VapA

6.1 Introduction

At the outer margins of bacterial cells are located different molecules whose functions are quite indispensable to their survival and natural history (Navarre and Schneewind, 1999). The architecture of bacterial envelopes is varied but essentially they fall into two groups with reference to the number of hydrophobic permeability barriers they possess; monoderms and diderms. The former are exemplified by the Gram-positive *Firmicutes*. Diderms possess an outer permeability barrier; these are varied in their composition and fine architecture having evolved different solutions to common functional challenges across varied lineages (Desvaux *et al*., 2009).

Following biosynthesis, proteins targeted to the cell envelope are translocated across the plasma membrane through a confined aqueous channel composed of sets of integral membrane proteins comprising the General secretory pathway (Sec) and the twin-arginine translocation pathway (TAT) (van Wely *et al.*, 2001). Some proteins are destined for an absolute secretion in which they operate at a distance to the cell. Those that are required to function in the outer margins but in a stable and defined position are subject to one of a number of modifications during or after this first secretion process. Although its primary function is to provide a rigid exoskeleton for protection against both mechanical and osmotic lysis, the cell wall of Gram-positive bacteria also serves as an attachment site for proteins that perform diverse tasks, especially mediating interaction of bacteria with their external environment which are often secured covalently to the peptidoglycan matrix *via* sortase enzymes (Navarre and Schneewind, 1999; Marraffini *et al.*, 2006). Due to the absence of a

retentive outer membrane, monoderms like the archetypal Gram-positive bacteria use *N*-terminal lipidation as another major mechanism by which proteins are retained in the periplasmic area; these lipoproteins can then be anchored *via* the modifying lipid to hydrophobic surfaces (Kovacs-Simon *et al.*, 2011).

Rhodococcus equi is a diderm; much research effort has been applied to its relative *Mycobacterium tuberculosis,* they share a highly conserved cell wall architecture. Its outer mycolate-rich lipid layer represents another potential anchoring point for lipoproteins that would allow their presentation at the cell surface (Kovacs-Simon *et al.*, 2011).

In addition to VapA being a secreted protein, it is also thought to be lipidated and surface-located (Tan *et al.*, 1995; Meijer and Prescott, 2004) though the nature of that lipidation is poorly understood. The lipid moiety could either be a permanent co-translational addition or a post-translational modification which may perform additional roles other than simple anchorage (Lavental *et al.*, 2010). The lipid modification of VapA is recognised as being unusual, partly from the fact that the post-processing *N*-terminus is not the cysteinyl residue expected of a common lipoprotein. Recently, channel-forming porins of *Corynebacterium glutamicum* (another mycolylated relative of *R. equi*) have been shown to bear *O*-linked mycolate residues (Huc *et al.*, 2010) and a mycolyltransferase (Cmt) has been identified to catalyse this transfer (Huc *et al*., 2013). Since lipoproteins and channel-forming porins are non-covalently attached to the wall, it is possible that a range of mycolylated proteins might be located at the outer membrane. Indeed, mycolylation might be an important intermediate step in the trafficking of a subset of secreted proteins to be retained at the outer surface of the mycolata.

Previous attempts to characterise the supposed lipidation of VapA *via* traditional methods have been unsuccessful. We sought to determine whether mycolylation of outer membrane proteins occurred in *R. equi* and then whether a molecular biology approach could address the hypothesis that VapA might bear a mycolyl residue.

6.2 Investigation of envelope protein-bound lipids of Rhodococcus equi

6.2.1 Phase separation of Rhodococcus equi membrane proteins

Membrane proteins were extracted from washed *R. equi* cells in buffer containing the non-ionic detergent Triton X-114. The detergent replaces most of the lipid molecules in contact with the hydrophobic units of the proteins producing a protein-detergent mixed micelle (Bordier, 1981). Hence *via* the amphiphilic nature of these proteins, they can be effectively separated from hydrophilic proteins which do not possess any hydrophobic moiety. Below 20°C, Triton X-114 molecules disperse in aqueous solutions into small micelles which form quite a homogenous solution but above 22°C, these micelles coalesce, becoming large enough to produce a phase separation (Kittelberger *et al.*, 1995; González de la Vara and Alfaro, 2009). Thus when Triton X-114 is used to extract membrane proteins and the temperature is raised above 22°C, hydrophobic proteins concentrate in the lower detergent phase as an oily droplet while the hydrophilic proteins separate in the upper detergent-free phase thus separation of hydrophobic and hydrophilic proteins can be achieved in a single step (González de la Vara and Alfaro, 2009).

As this aspect of the study is principally concerned with VapA, an experimental design that established the partitioning of this key protein during its fractionation was required. The VapA-specific monoclonal antibody Mab10G5 was ideal detection

reagent in this respect. After phase separation (as described in section 2.8.10), protein extracts from both aqueous and detergent phases were immobilised on a nitrocellulose membrane in a dot blot format and the presence of VapA was determined immunochemically using Mab10G5 as the primary antibody (Figure 6.1).

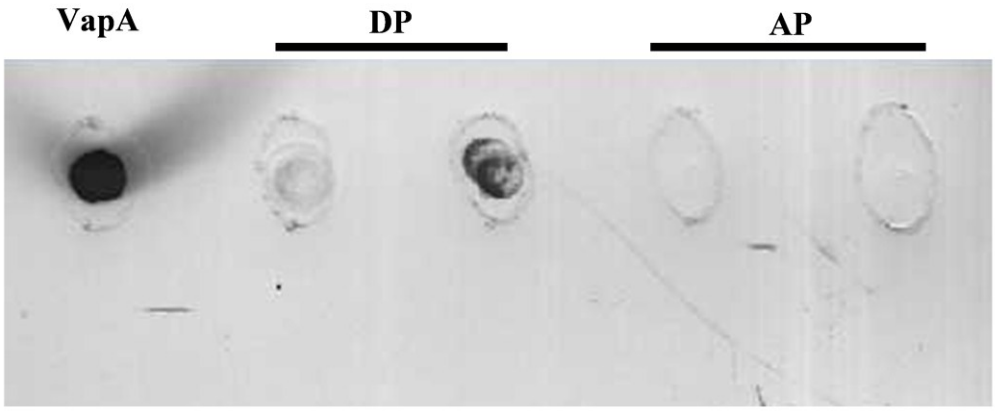


Figure 6.1 Fractionation of VapA in Triton X-114
Extracts of *Rhodococcus equi* 103^{+} membrane proteins were fractionated between Aqueous (AP) and Detergent (DP) phases. Samples were applied to nitrocellulose membrane in a dot blot format. The blot was probed with the VapA-specific monoclonal antibody Mab10G5 and visualised with an anti-mouse IgG (alkaline phosphatase conjugate): Samples: VapA (recombinant expressed in *E. coli* BL21 (DE3); DP, 2 x detergent phase extracts; 2 x AP, aqueous phase extract after phase separation. Two independent extractions are shown here.

A strong signal was seen for a recombinant VapA positive control confirming the validity of our stock antibody as a detection device. In all samples tested, a positive signal was detected within the detergent phase of *R. equi* extracts, although this varied in strength. This immunochemical reaction supported the detection of VapA in these cell wall-derived fractions. No signal was detected when analysing the aqueous phase in any experiment indicating a large partition coefficient (unmeasured) and lipophilicity for VapA. In the absence of structural information, it

was assumed this indicated that VapA was both present in the analyte sample and was probably lipidated. In order to formally identify VapA in this detergent phase, the proteins were trypsin-digested (in gel), separated by reverse-phase HPLC and subject to tandem mass spectrometry in which the fragment ASDTAGQEQQYDVHGDVISAVVYQR representing 13% coverage of the VapA amino acid sequence was detected (Figure 6.2).

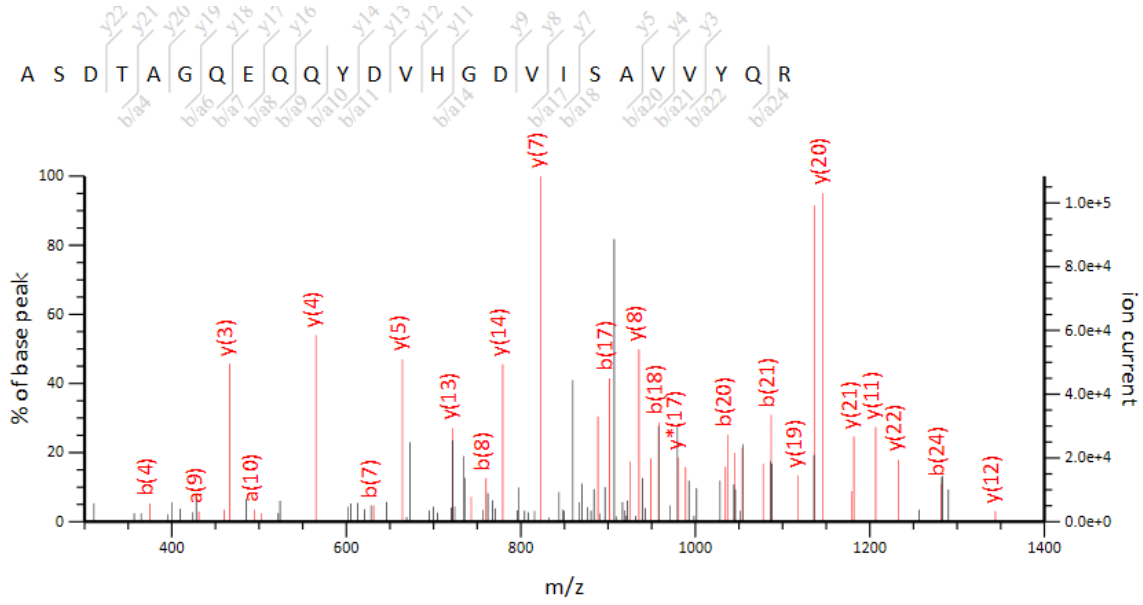


Figure 6.2 MS/MS fragmentation of ASDTAGQEQQYDVHGDVISAVVYQR

The fragment represents 13 % coverage of the amino acid sequence of VapA. Peptide score via Mascot is 56 %.

6.2.2 Lipid analysis

Virulence-associated proteins such as VapA may be tethered to the rhodococcal cell envelope by a lipid modification. With the emergence of mycoloylation as a post-translational modification of secreted proteins, it would be interesting to know if this

was also a feature in rhodococci. The site of mycolylation is known in only one case and as no region of strong homology exists with the other poorly-defined small mycolylated porin (Huc *et al.*, 2010), no important sequence motifs can yet be considered. *Rhodococcus equi* also contains porin proteins. As lipoproteins and channel-forming proteins are all non-covalently attached to the envelope, several proteins of *R. equi* might be bound to the wall *via* a mycolyl linkage. The working hypothesis is that proteins destined to be surface-exposed are mycolylated during secretion and consequently targeted to the mycolate-based outer membrane.

To perform an analysis of protein–bound lipid within the fractions produced by Triton X-114 phase separation, proteins were precipitated using trichloroacetic acid, lyophilised overnight, and re-suspended in a methanol:NaCl solution.

A major concern was that the extract may include mycolic acid-containing glycolipids that could contaminate the protein preparations and lead to a false interpretation that mycolic acid methyl esters had been derived from proteins. The solution was first delipidated by phase partition by mixing with petroleum ether (as described in section 2.9.1.1) to remove free lipids including the mycolate-containing glycolipids like the predominant trehalose (mono-/di-) mycolate. In order to understand the complexity of the lipid mixture removed at this point, the upper phase containing the free lipids was analysed by TLC (figure 6.3.).

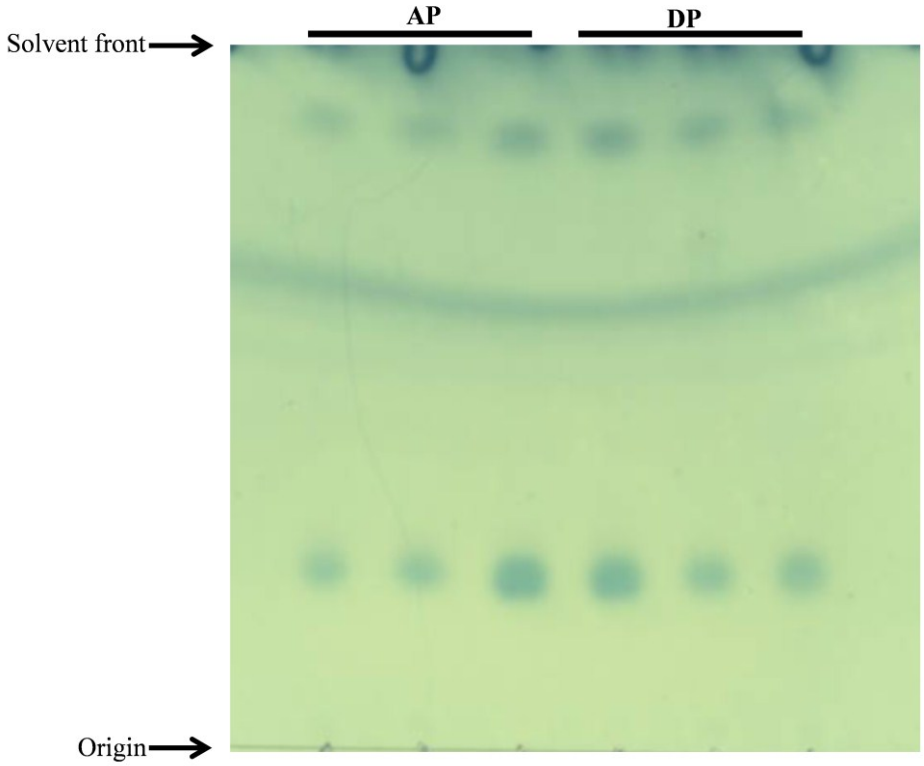


Figure 6.3 TLC Analysis of free lipids from *R. equi* subjected to Triton X-114 phase partitioning on silica.
AP, aqueous phase extracts; DP, detergent phase extracts. Five microlitre samples from equivalent extracts were loaded, the TLC was developed using chloroform:methanol (2:1). Lipids were visualised by charring with 5% ethanolic phosphomolybdic acid.

When the plate was stained with molybdophosphoric acid and visualised, major bands were detected migrating with R_f = 0.2 and 0.8 and there was no significant difference in band pattern observed between lipids in the detergent phase and the aqueous phase extracts (Figure 6.3). These materials have not been further characterised but are likely to be phospholipids and free fatty acids respectively. There appeared to be no significant population of glycolipids in this extract since no bands were observed when an identical plate was visualised after staining with the carbohydrate-reactive 1-Naphthol (Results not shown).

The lower methanolic phase containing extracted protein was lyophilised and treated with 5 % aqueous tetrabutylammonium hydroxide at 100°C to release protein-bound lipids, which were then converted to methyl esters (as described in section 2.9.1.3). A similar process was carried out for proteins precipitated from the original aqueous phase of the Triton X-114 partitioning.

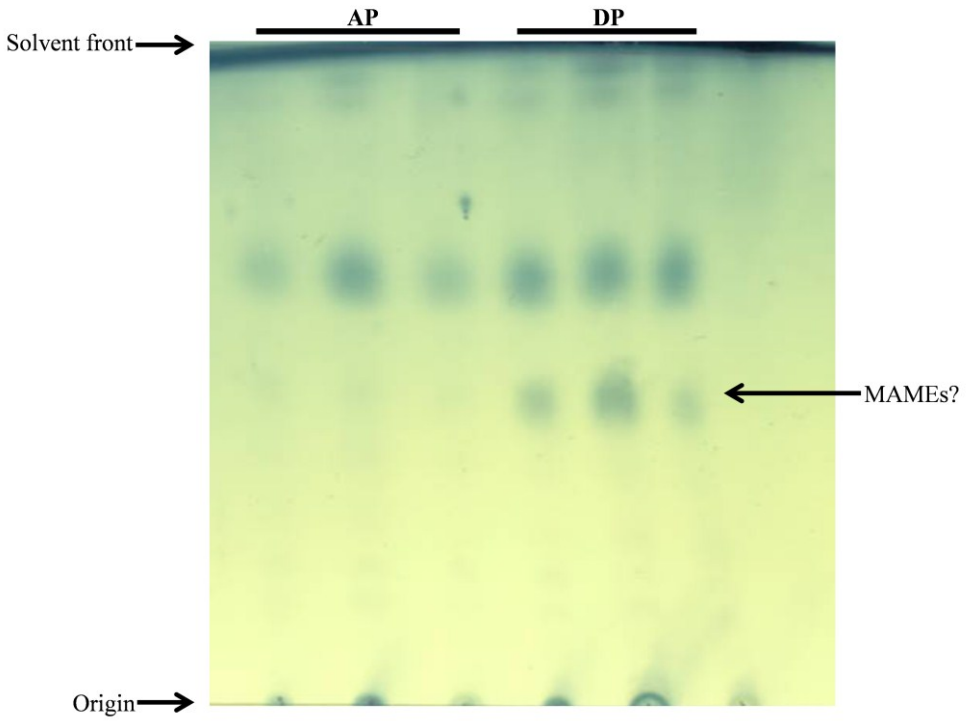


Figure 6.4 Analysis of acyl-methyl esters released from protein-bound lipids from *R. equi* cells subjected to Triton X-114 phase partitioning
AP, aqueous phase extracts; DP detergent phase extracts. MAMEs?, putative mycolic acid methyl esters. Five microlitre samples from equivalent extracts were loaded, the TLC was developed using petroleum ether: acetone (95:5). Lipids were visualised by charring with 5% ethanolic phosphomolybdic acid.

TLC analysis of methyl esters (Figure 6.4) revealed significant quantities of a common material both in the aqueous phase and the delipidated detergent phase extracts ($R_f = 0.66$) (these varied markedly in relative intensity with each extract, see

Figure 6.6) but methyl esters migrating at R_f = 0.46 only appear in the detergent phase extract. Because of their relative migration in this solvent system, these lipids at R_f = 0.46 were believed to be mycolic acid methyl esters (MAMEs?). This indicates that hydrophobic proteins tethered in the wall of *R. equi* could possess mycolyl residues and thus a mycolylation of VapA could be considered. Lipid samples were later subjected to MALDI-TOF MS to confirm the identity of the lipids derivatised as methyl esters (see section 2.9.4 and Figure 6.7).

In order to increase confidence in our interpretation of this interesting finding, the protein-bound lipids (derivatised as methyl esters) of *Rhodococcus equi ΔkasA* were also investigated in the same way. *Rhodococcus equi* Δ*kasA* is modified *via* a deletion of the gene for β-ketoacyl-(acyl carrier protein)-synthase A (KasA), the key fatty acyl chain elongation activity of the FAS-II system responsible for synthesis of the long-chain meromycolic acid component of the mycolic acids. Thus the strain produces mycolic acids that are about ten carbons shorter in chain length than those of the parental strain (Sydor *et al.*, 2013, see figure 6.5) suggesting condensation of FAS-I products to form short chain mycolates only.

The extraction of protein-bound lipids from *R. equi*Δ*kasA* revealed a slow-migrating acyl methyl ester on TLC. This material migrated to a similar R_f to that of the putative MAMEs recovered from the parental strain but consistently migrated to a lower R_f (0.32) (Figure 6.6). This chromatographic behaviour is consistent with an increased polarity related to the loss of acyl chain length in the mutant extract and consequently greater interaction of the polar β-hydroxy and residual terminal carboxy- moieties with the silica matrix.

Taken together, these findings are wholly consistent with the demonstration of protein mycolylation in the outer cell wall of *R. equi*.

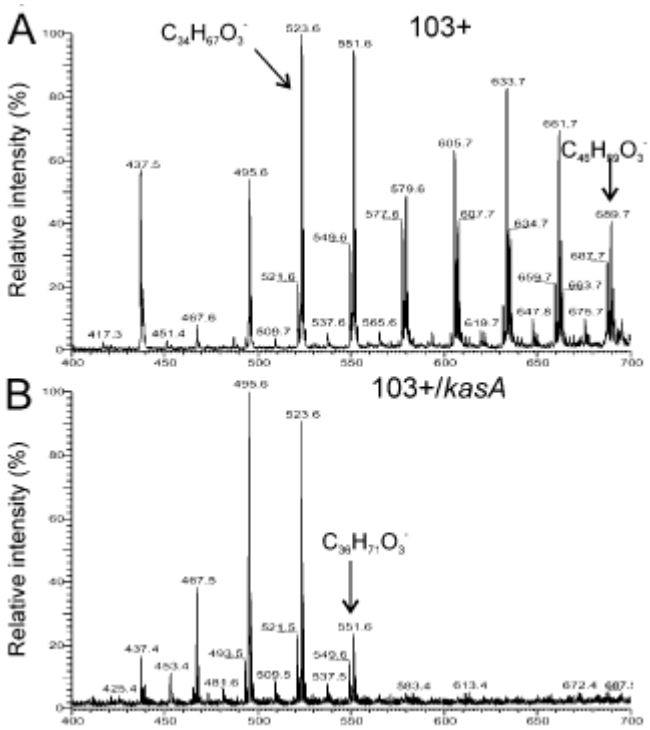


Figure 6.5 Mass spectrometry (with electrospray ionisation) (as m/s) of mycolic acids from chloroform:methanol extracts of *R. equi*
(A) wild type, from (B) *kasA* mutant (from Sydor et al., 2013). Their profiles show that many mycolic acids found beyond m/z 551.6 in the wild type seem to be lost in the mutant (Adapted from Sydor *et al.*, 2013).

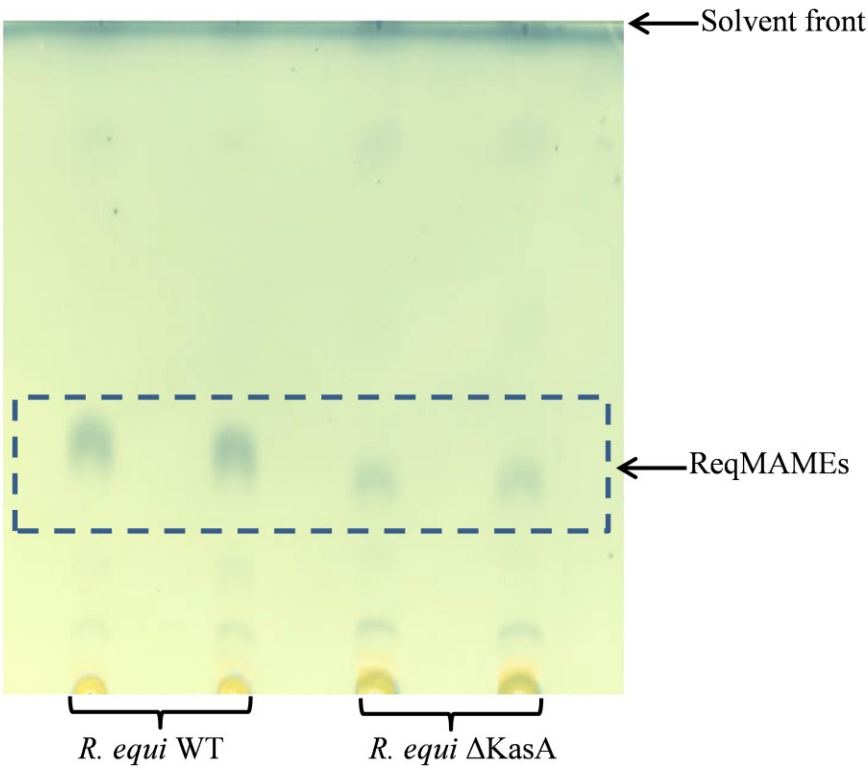


Figure 6.6 TLC of mycolic acid methyl esters of *R. equi*.
Methyl esters were released from the protein-bound lipids of the wild type (WT) and Δ*kasA* mutant strains of *Rhodococcus equi* subjected to Triton X-114 phase partitioning. Five microlitre samples from equivalent extracts were loaded, the TLC was developed using petroleum ether: acetone (95:5). Lipids were visualised by charring with 5% ethanolic phosphomolybdic acid.

6.2.3 Identification of acyl-methyl esters via MALDI-TOF MS

The TLC profiles in figures 6.4 and 6.6 appear to show that some of the acyl methyl esters detected in this study could be putative mycolic acid methyl esters. In order to confirm the identities of these lipids, the samples were subjected to MALDI-TOF MS.

Following Triton X-114 phase partitioning, protein-bound lipids (from *R. equi* 103S and *R. equi ΔKasA*) were extracted from the detergent phase and derivatised as

methyl esters as described earlier. Using α-Cyano-4-hydroxycinnamic acid as the matrix, MALDI-TOF MS (Figure 6.7) mass profiles mirror those seen from the corresponding crude cell wall extracts of the parental and *kasA* mutant strains (see figure 6.5, Sydor *et al*., 2013). Both strains produced similar MAMEs up to m/z 568. The maximum chain length of the protein-bound lipids from *R.equi* Δ*kasA* was 36 carbons (m/z 568, $C_{36:0}$Me) while the wild type strain produced significant amount of lipids beyond m/z 568 (figure 6.7, Table 6.1). For the mutant, residual signals for these larger lipids seem to be close to background levels.

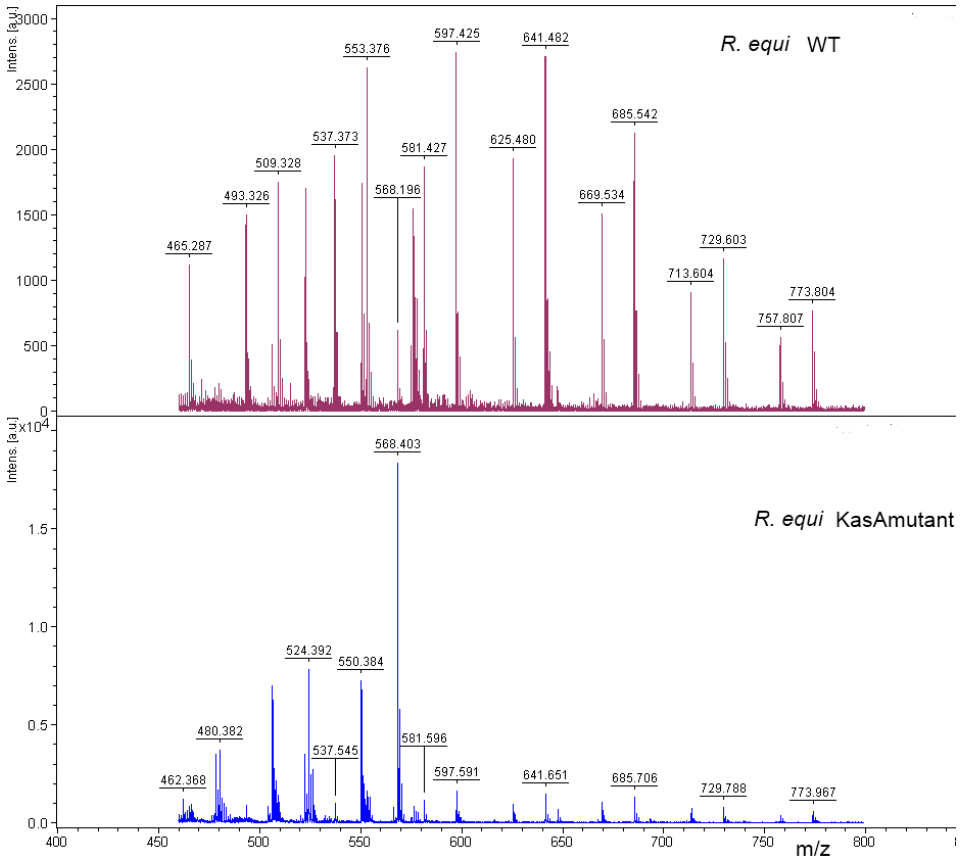


Figure 6.7 MALDI-TOF spectra of mycoloyl acid methyl esters released from protein-bound lipids from the *R. equi* 103S and ΔkasA strains

Taken together the combined results of the survey of protein-bound lipids in the walls of *R. equi* and its isogenic mutant in *kasA* were wholly consistent with mycolylation of proteins in this species. Immunochemical and mass spectrometric techniques established that VapA was among the proteins present in these extracts, supporting the hypothesis that lipid modification of this key virulence determinant could be *via* mycolylation.

However, in the light of structural information identifying the relatively apolar surface of the core structure of the VapA model (Figure 4.10), a retrospective analysis of the partitioning of recombinant VapA from *E. coli,* and therefore unlikely

to be lipid-modified, revealed that this molecule partitioned to the detergent phase in triton X-114. This might indicate that the 'top' of the protein structure has enough hydrophobic nature to maintain an association with the mycolate membrane independently of a lipid modification.

Table 6.1 Protein-bound lipids predominant in *R. equi* WT revealed by MALDI-TOF MS

	Modelled configuration of putative mycolic acids	
m/z	Number of Carbons (+ Me*)	Number of unsaturations
581.4	37	0
597.4	38	0
625.5	40	0
641.5	42	4
669.5	44	4
685.5	45	3
713.6	47	3
729.6	48	2
757.8	50	2
773.8	51	1

*Species were derivatised as methyl esters by treatment with iodomethane

6.3 Lipidation of Recombinant VapA in Corynebacterium glutamicum

6.3.1 Cloning of vapA into the inducible E. coli-C. glutamicum shuttle vector pEKEx2

At the outset of this project, a novel post-translational modification was recognised in *C. glutamicum*, when it was determined that two porins are *O*-mycoloylated (Huc *et al.*, 2010). If such a function were to be conserved across the mycolata, this could

provide a potential explanation for the unusual lipidation of VapA and suggest a generic means for partitioning proteins specifically to the outer mycolic acid containing membranes of the mycolata. Due to the similar cell wall morphology shared by these two organisms, it was logical to determine whether expression of *vapA* in *C. glutamicum* would result in its mycoloylation. Furthermore, the subcellular localisation of VapA in this recombinant system might represent a useful means to develop an assay as a proxy for lipid modification. In this way a panel of site-directed mutants might be readily screened to suggest a site(s) for the lipidation.

Sequences encoding full-length VapA (designated *CgvapA* for expression in *C. glutamicum*) and its hexa-histidine tagged variant (designated *CgvapAhistag*) were amplified *via* PCR using oligonucleotide primers compatible with cloning in the *E. coli*-*C. glutamicum* shuttle vector pEKEx2 and analysed *via* agarose gel electrophoresis (Figure 6.8). *CgvapAHistag* was designed in order to introduce a *C*-terminal hexa-histidine tag in the protein product to facilitate (if need be) its purification *via* affinity chromatography or detection in immunoblots. As pEKEx2 has no coding sequence for hexahistidine tag, pET23-*vapA*full, which contains the entire *vapA* sequence, was used as the template for PCR to provide the code.

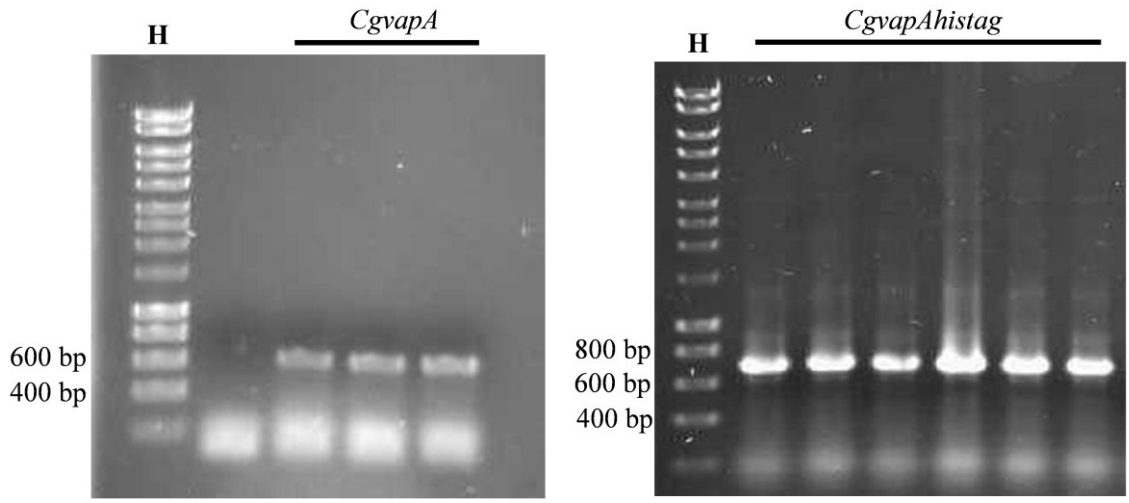


Figure 6.8 PCR amplification of *CgvapA* and *CgvapAhistag*
Lane H, 10 kbp size standard

The DNA band that migrated similarly to the 600 bp marker was consistent with expectation for *vapA* (570 bp). Similarly, the product destined for the histidine tagging migrated between between the 600 bp and 800 bp marker bands was consistent with the expected 727 bp *vapAhistag* amplicons (Figure 6.8). Samples that contained amplified products were pooled together and purified from agarose after electrophoresis. Each was subjected to restriction digestion using the enzymes *Kpn*I and *EcoR*I. The shuttle vector pEKEx2 was also digested using the same protocol and ligated to the digested PCR products. Recombinant plasmid DNA was later purified from transformants of *E. coli* Top 10. Analysis of the recombinant plasmids *via* agarose gel electrophoresis after restriction digestion is shown in figure 6.9.

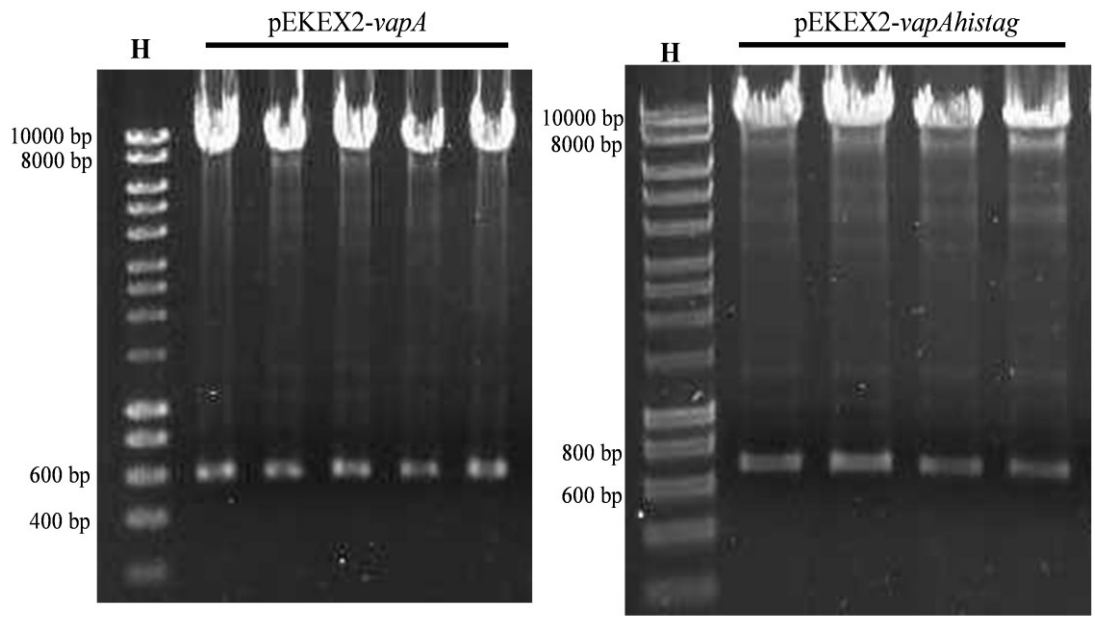


Figure 6.9 Restriction digestion of pEKEx2-*vapA*; and pEKEx2-*vapAhistag* with KpnI and EcoRI
Lane H is 10 kbp size standard marker (hyperladder 1).

The pattern of the visible bands was compatible with expectations strongly indicating that *vapA* and *vapAhistag* were effectively ligated to pEKEx2, this was validated by nucleotide sequence analysis.

6.3.2 Expression and subcellular localisation of recombinant VapA in Corynebacterium glutamicum

These pEKEx2-derived plasmids were used to transform electrocompetent *Corynebacterium glutamicum* ATCC 13032. Thereafter, antibiotic-resistant colonies were grown and subjected to IPTG induction of gene expression. Cells were harvested and protein isolated from both pellet and the spent culture supernatant. Initial analyses of subcellular location focussed on the non-particulate fractions, *i.e.* a clarified lysate and secreted proteins, for ease of processing. It was anticipated that a putative mycolylated VapA state might not be replicated in a heterologous host and

that a non-modified VapA might be secreted to the culture medium. Western blot analysis (probed with Mab10G5) revealed the presence of heavy bands migrating at

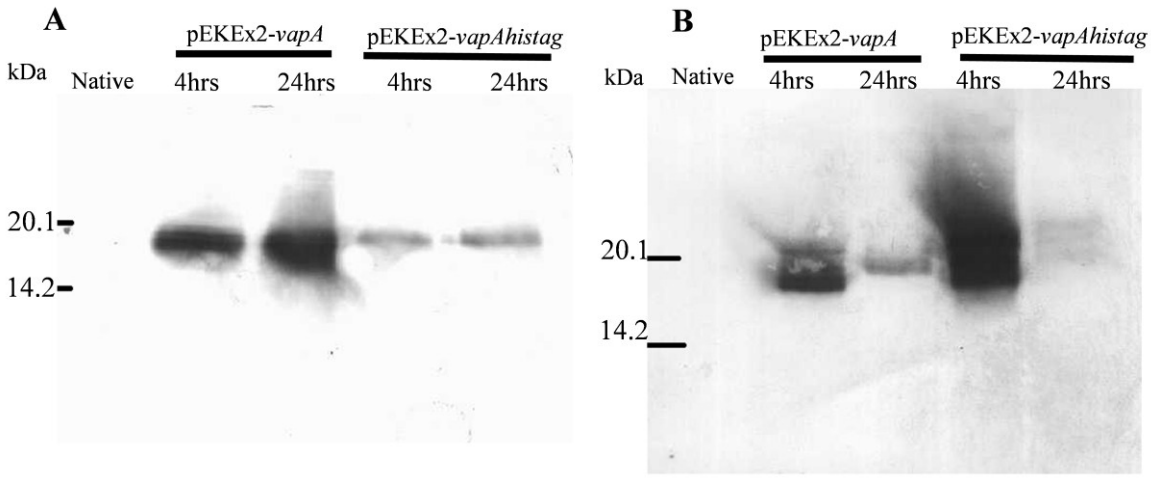


Figure 6.10 Western blot analysis of *C. glutamicum* proteins

(A) Clarified lysate from cell pellet (B) proteins from secreted protein fraction. Figure shows blot from native *C. glutamicum* (control); and *C. glutamicum* transformed with either pEKEx2-*vapA* or pEKEx2-*vapAhistag* and harvested 4 hrs post induction; or 24 hrs post induction.

a rate between those of the 14 kDa and 20 kDa standards in both the clarified lysate and broth secreted protein fractions which were consistent with the expected but unusual electrophoretic migration of native *R. equi* VapA, that was not well reproduced when *vapA* was expressed in *E. coli*. The diffuse band observed is in concord with the reported electrophoretic behaviour of VapA (which has been reported to span 18 – 22 kDa in some cases) (Tan *et al.*, 1995). This diffuse band was not detected in fractions from the non-transformed *C. glutamicum* cells inferring that these immunochemical reactions indicated *bona fide* detection of recombinant VapA.

It was clear however that this overproduced recombinant VapA was not exclusively located at the surface of the cell. Detection of VapA in the clarified lysate

(representing cytosolic and plasma membrane locations) could be anticipated through translation and secretory processing. Detection of significant quantities of recombinant protein in a cell-free fraction was anticipated as an indicator of poor lipid modification. This observation could indicate that either recombinant VapA was not modified and retained at the cell surface in *C. glutamicum* at all or that this process was not well regulated and secreted and modified populations of the recombinant protein might co-exist.

In order to rigorously identify recombinant VapA in these fractions, they were subjected to SDS-PAGE, protein bands (already stained corresponding to the molecular size of interest) were excised and subjected to in-gel trypsin digestion and LC-tandem MS. Raw chromatographic files were processed into Mascot compatibility files and searches performed using the online Mascot MS/MS ions search function against the NCBI non-redundant protein database with taxonomy set at Actinomyces (Table 6.2). The search parameters applied were peptide mass tolerance of ± 0.5, fragmented mass tolerance of ± 0.5, peptide charge of 2+ and 3+ and semitryptic specificity allowing for up to two missed cleavages. Other parameters were fixed modifications at carboxymethyl (of cysteine) and variable modifications with deamidation at asparagine and glutamine. Details of the score, number of non-redundant peptides identified, and the amino acid coverage of VapA sequence (NCBI Reference Sequence: NP_858467.1) are shown in table 6.2. LC-MS/MS showed that peptides for VapA amino acid sequence were effectively identified both in the clarified lysate (CLY) and the broth supernatant proteins (BSP) with significant scores. All the fragments identified exceeded Mascot's significant

level (Mascot score > 40). Thus it could be concluded that VapA was effectively expressed in *C. glutamicum* to a significant degree.

Table 6.2 Confirmation of recombinant VapA from various fractions of *C. glutamicum* expressing recombinant vapA via LC-MS/MS.

Fraction	Induction period	Score	No of Matched peptides	Protein coverage	Non-redundant peptides
C. glutamicum* pEKEx2-*vapA					
CLY	4	67	1	13	ASDTAGQEQQYDVHGDVISAVVYQR
CLY	24	75	1	13	ASDTAGQEQQYDVHGDVISAVVYQR
CLY	4	67	1	13	ASDTAGQEQQYDVHGDVISAVVYQR
SP	24	63	1	13	ASDTAGQEQQYDVHGDVISAVVYQR
C. glutamicum* pEKEx2-*vapAhistag					
CLY	4	138	4	17	FHVFGPEGK ASDTAGQEQQYDVHGDVISAVVYQR ASDTAGQEQQYDVHGDVISAVVYQR ASDTAGQEQQYDVHGDVISAVVYQR
CLY	24	120	2	17	FHVFGPEGK ASDTAGQEQQYDVHGDVISAVVYQR
SP	4	127	3	17	FHVFGPEGK ASDTAGQEQQYDVHGDVISAVVYQR ASDTAGQEQQYDVHGDVISAVVYQR
Position of detected fragments in native VapA sequence					
MKTLHKTVSKAIAATAVAAAAAMIPAGVANATVLDSGSSSAILNSGAGSGIVGSGSYDSS TTSLNLQKDEPNGRASDTAGQEQQYDVHGDVISAVVYQRFHVFGPEGKVFDGDAGGLTLP GAGAFWGTLFTNDLQRLYKDTVSFQYNAVGPYLNINFFDSSGSFLGHIQSGGVSTVVGVG GGSGSWHNA					

CLY indicates samples derived from a clarified lysate; SP indicates samples derived from secreted proteins and Q indicates a deamidated glutamine residue. Text in Grey indicates that cleaved signal peptide of VapA. Highlighted text identifies detected peptides.

A few glutamine residues on some of the redundant peptides are evidently deamidated. Non-enzymatic deamidation of glutamine and asparagine is a post-

translational modification that accompanies *in vitro* purification, storage and handling of proteins initiated by acids and the presence of non-bulky amino acids near glutamine in the protein's primary sequence (Volkin *et al.*, 1997; Joshi *et al.*, 2005). Thus trichloroacetic acid (though a weak acid) used in protein precipitation maybe responsible for the deamidation of glutamine in the tryptic fragments of VapA in the broth supernatant proteins. Though several of the peptides have glycinyl/alaninyl residues neighbouring the deamidated residue, secondary structure effects appear a less likely cause (Joshi *et al.*, 2005).

6.4 Analysis for corynomycolate modification and compartmentalisation of recombinant VapA in C. glutamicum

Protein extracts following *vapA* expression in *C. glutamicum* were delipidated and protein-bound lipids extracted and analysed by TLC as before (Figure 6.11). TLC analysis of the preparation containing acyl methyl esters shows a common band (R_f = 0.5, labelled **P**) in the clarified lysate and the secreted protein fraction of the transformants and the native *C. glutamicum* (Figure 6.11). However, the abundance of another common band (which occurs at R_f = 0.3, indicated as CgMAMEs) is significantly increased in the cells transformed with pEKEx2-*vapAhistag* (harvested 4 hrs and 24 hrs post induction) and pEKEx2-*vapA* harvested 24 hrs after IPTG induction (Figure 6.11A) in both the clarified lysate and more markedly in the secreted protein fraction (Figure 6.11 B). This indicates that there is a pronounced increase in this unusual protein lipidation in *C. glutamicum* coincident with the over expression of *vapA*. These are described as putative MAMEs due to the similarity of their migration to *C. glutamicum* MAMEs described elsewhere (Gande *et al.*, 2004).

Extracts of *C. glutamicum* pEKEx2-*vapA* harvested 4 hrs after induction with IPTG seem to show no difference in lipid pattern when compared with native cells (lane 1).

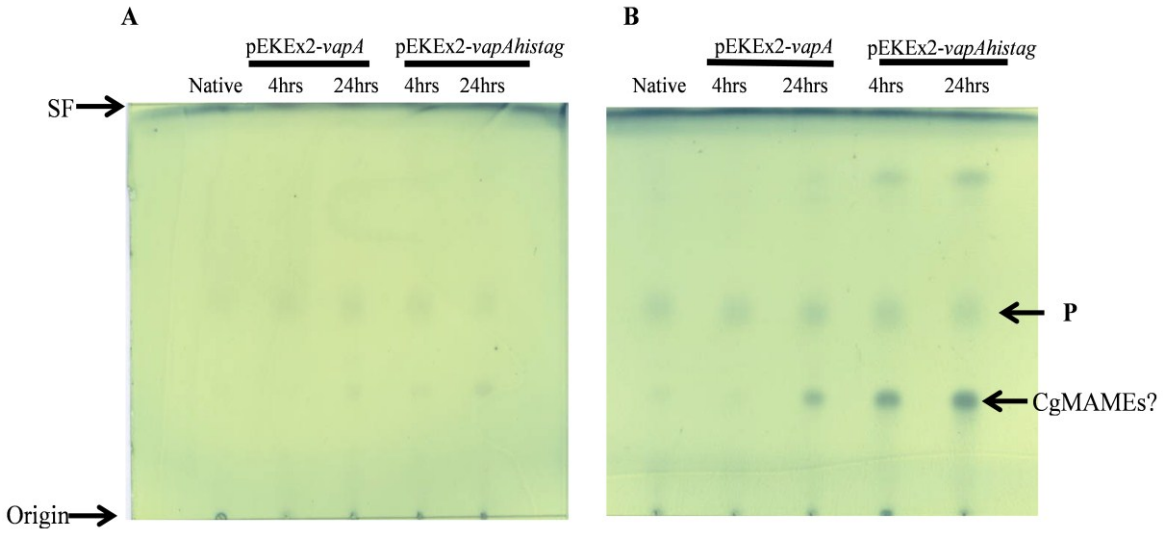


Figure 6.11 TLC analysis of protein-bound lipids of delipidated extracts
(A) Clarified lysate (B) secreted protein fraction. Figure shows profile from native *C. glutamicum* (control); and *C. glutamicum* transformed with either pEKEx2-*vapA* or pEKEx2-*vapAhistag* and harvested 4 hrs post induction; or 24 hrs post induction. CgMAMEs?, putative corynomycolic acid methyl esters. SF, solvent front. Five microlitre samples from equivalent extracts were loaded, the TLC was developed using petroleum ether: acetone (95:5). Lipids were visualised by charring with 5% ethanolic phosphomolybdic acid.

The consistency of abundance in 'P' in these extracts validates the equivalent loading of the TLC samples. This apparent poor yield of MAMEs may indicate that the degree of post translational lipidation may occur more slowly with this construct as these bands became significantly pronounced in samples harvested 24 hrs post induction (figure 6.11). It is possible that this lipidation profile might be secondary to a limiting polypeptide yield at this time point (not established).

The accumulation of protein-bound mycolic acids in the spent medium was not expected as accumulation of modified VapA in the outer membrane was envisaged. It is possible that accumulation of the protein at this site may have led to the formation of protein-bearing outer membrane vesicles. These are spherical, bilayered proteolipids with an average diameter of 20–200 nm that encapsulate various compounds such as lipopolysaccharide, lipids, soluble or membrane-associated proteins, genetic materials, and other factors associated with virulence in both Gram-positive and Gram-negative bacteria. They perform various functions such as the transfer of proteins, cell-to-cell signalling, the elimination of competing organisms, and the delivery of toxins to host cells and are conserved across Gram-negative bacteria though they have also been reported in Gram-positive organisms such as *Staphylococcus aureus* (Lee *et al.*, 2009; Berleman and Auer, 2013). They are discrete, closed outer membrane blebs produced by growing cells and not products of cell lysis or cell death which can often be deposited from a cell-free supernatant (Kuehn and Kesty, 2005). In *C. glutamicum*, the shedding of significant amount of membrane vesicles and/or membrane fragments to the external culture medium is reported to be dependent on a highly perturbed cell wall which may occur as a result of deficiency in cell wall synthesis (Raad *et al.*, 2010). In this case, accumulation of recombinant VapA might be envisaged to cause disruption to the structure of the outer membrane. However when the culture supernatant was centrifuged at 100,000 x *g*, no discernible pellet was formed, thus there was no visible evidence that outer membrane vesicles were shed. Performing the experiment at a larger scale (to increase the likelihood of detection) also showed a similar pattern thus the accumulation of VapA in the membrane may not cause significant disruption to the integrity of the cell wall. However, it is also possible that the density of such blebs

might have caused them to be buoyant and that a thin layer of lipid material could have been overlooked at the surface.

The next issue addressed was whether recombinant VapA was retained in the corynebacterial cell wall. Transformants (with pEKEx2-*vapA*) were grown and subjected to IPTG induction overnight. Proteins were extracted from both a plasma membrane fraction and the debris after clarification of the extract (which contains deposited wall materials) with SDS-Tris buffer and Urea-Tris buffer (as described in sections 2.8.13). Protein extracts were later analysed by Western blot using Mab10G5 as the primary antibody. These immunoblots show that VapA was detected in proteins from both the debris and the plasma membrane using both extraction systems (Figure 6.12). The retention in the cellular debris might suggest appropriate mycoloylation and location in the outer membrane of the heterologous host.

The detection of a strong signal for VapA in the plasma membrane might indicate two populations of recombinant protein retained *via* different means. Firstly, some of the recombinant VapA might be associated with this fraction transiently, having been partially secreted at the point of extraction. Another explanation is that there is a contribution from mycoloylated protein retained here as an artefact of over expression and inappropriate/incomplete partitioning to the mycoloyl layer. Neither scenario has been rigorously investigated.

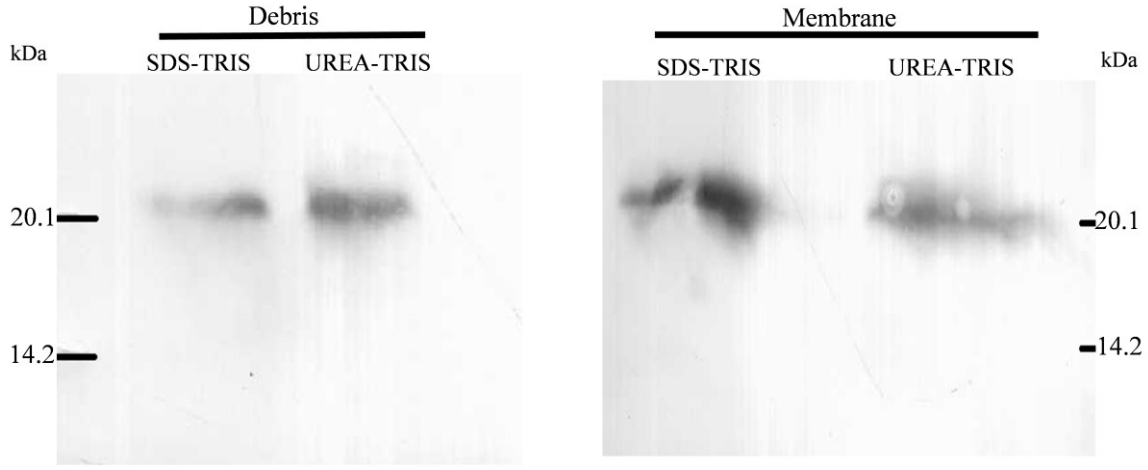


Figure 6.12 Retention of recombinant VapA in cell wall material.
Proteins were extracted using either SDS-Tris buffer or Urea-Tris buffer plasma membrane or wall-rich debris fractions. Proteins were separated using SDS-PAGE and immobilised on nitrocellulose membrane. VapA was visualised using Mab10G5 as a primary antibody.

This preliminary result seemed to show that recombinant VapA could be bound to the cell wall of *C. glutamicum* in addition to being secreted. In order to improve our understanding of the subcellular distribution of VapA, the experiments were refined; the cell debris was washed and then subjected to density gradient centrifugation after re-suspension in 60 % Percoll. The final buoyant cell wall layer was washed free of Percoll in PBS. The clarified lysate was also subjected to ultracentrifugation to deposit the plasma membrane fraction as described in section 2.8.14. After standardising the total protein content of each, these plasma membrane and cell wall fractions as well as a secreted protein fraction (the cell-free culture supernatant) were analysed by Western blotting using MAb10G5 (Figure 6.13).

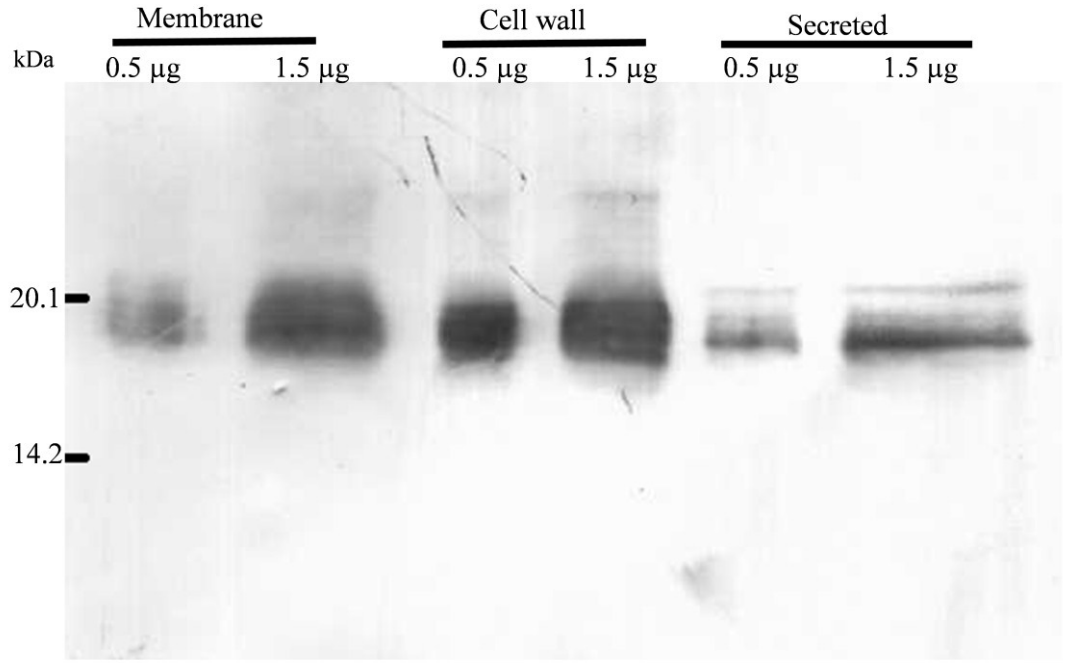


Figure 6.13 Refined fractionation of Recombinant VapA expressed in *C. glutamicum*.
Proteins were extracted from the plasma membrane and purified cell wall fractions using SDS. Protein content-standardised samples of these fractions were electrophoresed alongside a secreted protein fraction from the same culture and immobilised on nitrocellulose membrane. VapA was visualised using Mab10G5 as a primary antibody.

This enhanced protocol for following the subcellular compartmentation of VapA showed that the virulence-associated protein is primarily retained in the cell wall fraction of *C. glutamicum*, is considerably represented in the plasma membrane and is also secreted (Figure 6.13). Thus, the processing of VapA could be similar in both *R. equi* and *C. glutamicum*.

6.5 Site-directed mutagenesis

Having established that protein extracts that include VapA from *R. equi* contain apparently covalently-linked mycolic acids, and that the over expression of

recombinant *vapA* in *C. glutamicum* causes an apparent marked mislocalisation of significant quantities of corynomycolic acids, and the protein is retained in purified cell wall material, we were encouraged to develop an alanine-scanning mutagenesis study to identify a lipidation sites using cell wall retention as a measure of lipidation.

Mature VapA does not possess the terminal acylated *N*-terminal cysteine motif characteristic of membrane-anchored lipoproteins (Tan *et al.*, 1995); in fact preproVapA does not even contain a cysteine residue. However, analysis of preproVapA *via* SignalP predicted its likely cleavage by signal peptidase between Ala31 and Thr32 (see Appendix E). A simply-drawn hypothesis is that this *N*-terminal threonine residue might be the site of an *O*-linked mycolylation. Countering this hypothesis is the evidence that *N*-terminal sequencing of mature native VapA has been achieved identifying Thr32 as the first product of the Edman degradation process (IC Sutcliffe, unpublished results). Likewise, if it is expected that VapA and VapB are similarly processed, Thr32 is not conserved.

However, a Thr32Ala mutation was sought in the pEKEx2-*vapA* construct to test the hypothesis. Threonine has a hydroxyl side chain which could participate in *O*-linked modifications thus could provide a means of post-translational modifications of proteins (Mukherjee *et al.*, 2007). Hence the substitution of threonine with alanine removes that possibility because of the lack of a hydroxyl group in alanine. Oligonucleotide primers were first designed for the amplification of pEKEx2-*vapA* in order to introduce a point mutation in *vapA*. The aim was to substitute Thr32 to Ala by changing the codon ACC to GCC (using the primers shown in table 2.5). The plasmid pEKEx2-*vapA* was used as the template and amplified *via* PCR. After PCR, the mixture was subjected to restriction digestion using *Dpn*I in order to digest *dam*

methylated parental DNA originally obtained from the *dam* positive *E. coli* Top10. Figure 6.14 shows an agarose gel electrophoresis image of the amplified plasmid before and after *Dpn*I treatment.

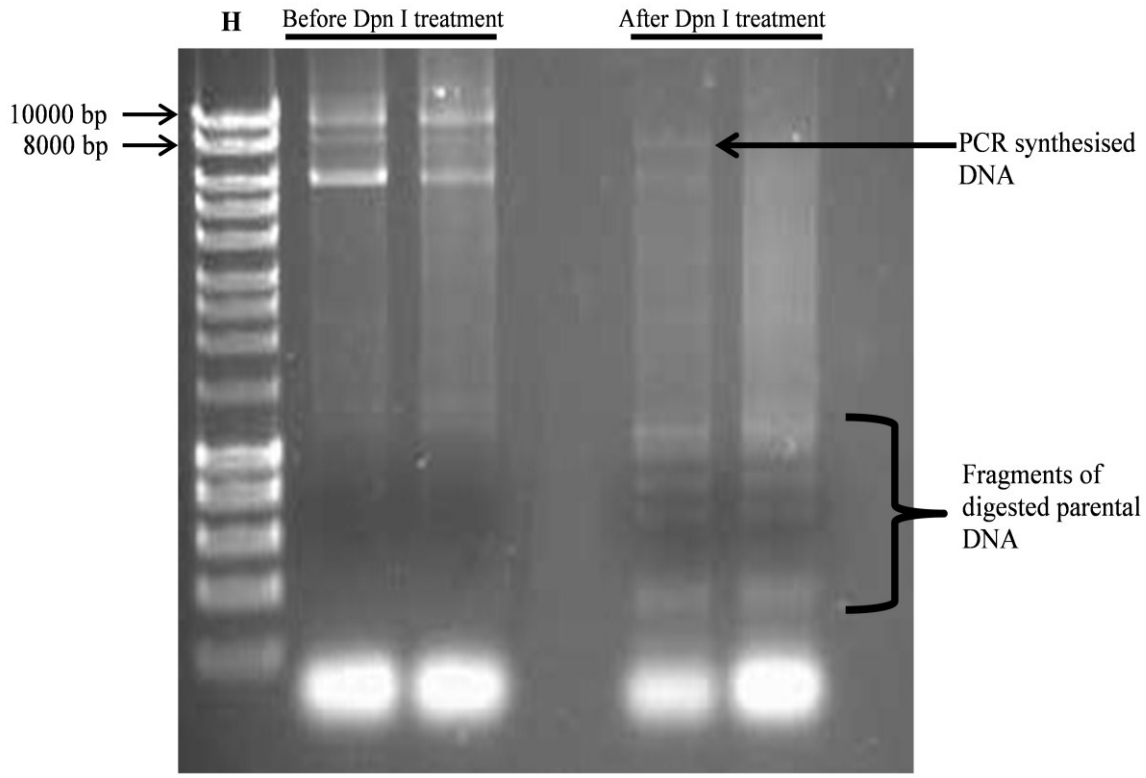


Figure 6.14 Gel image after amplification of pEKEx2-*vapAmutant*

Figure shows products before and after DpnI treatment. Treatment of the product following PCR is designed to eliminate parental DNA (template DNA).

The recombinant plasmid (PCR synthesised DNA) was used to transform competent *E. coli* Top10 cells and plasmid DNA was extracted after overnight culture of antibiotic-resistant colonies. In order to ensure that *vapA*Thr32A was present, the plasmid was subjected to restriction digestion analysis using the enzymes *EcoR*I and *Kpn*I.

The fluorescing bands pattern (data not shown) confirms the migration of pEKEx2 (about 8161 bp) and *vapA*Thr32Ala (570 bp) thus indicating the presence of the insert DNA in the plasmid. The identity and integrity of the insert and mutation was confirmed by nucleotide sequence analysis. *Corynebacterium glutamicum* was later transformed with the plasmid to study the impact of the mutation on the subcelluar compartmentation of VapAmutant in that organism.

6.6 Subcellular compartmentation of VapAmutant in Corynebacterium glutamicum

Investigation of the behaviour of the Thr32Ala mutant in *C. glutamicum* revealed that the protein was still retained in the membrane and attached to the cell wall indicating that the mutation may not have a significant effect on the localisation of the protein (Figure 6.15) and perhaps lipidation. This may indicate that Thr32 is not the site for VapA lipidation or that this single point mutation did not ablate functional lipidation, *i.e.* multiple sites might be lipidated.

Since substitution of Thr32 to Ala seemed not to affect the subcellular localisation of VapA in *Corynebacterium glutamicum*, other mutants were sought since VapA possesses many other potential sites for *O*-acylation (Figure 6.16). In order to rapidly screen the entire Ser/Thr population in VapA, a strategy was developed that took advantage of clusters of these residues. These clusters were mutated as blocks in which all residues in that locale were mutated simultaeously. Should a profound change to subcellular location be evident, single mutations would be explored in each case to improve the resolution of the strategy to the single residue level.

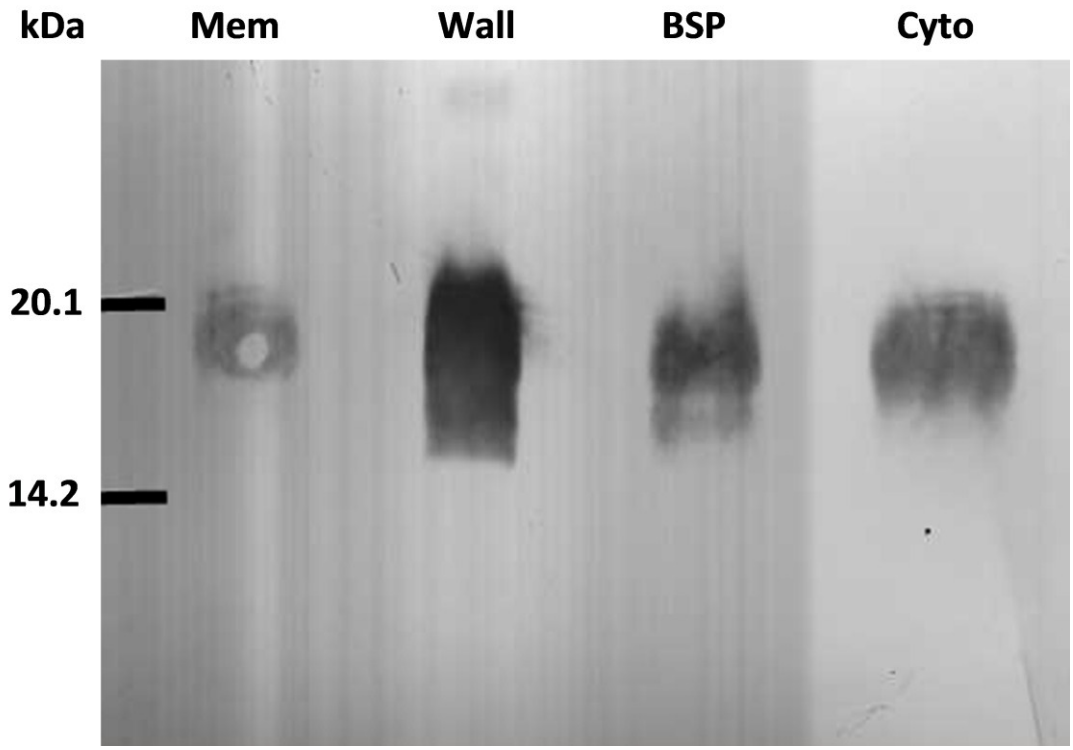


Figure 6.15 The Thr32Ala mutation does not perturb recombinant VapA localisation in *C. glutamicum*

Fractions of *C. glutamicum* pEKEX2-*vapA*Thr32Ala were electrophoresed and immobilised on nitrocellulose membrane. VapA was detected using Mab10 G5 as primary antibody. Fractions: Mem, plasma mebrane; Wall, purfied cell wall; BSP, secreted protein; Cyto, cytosol . Standardised loading of 5 µg protein were used throughout.

```
  1   MKTLHKTVSK AIAATAVAAA AAMIPAGVAN ATVLDSGSSS AILNSGAGSG IVGSGSYDSS
 61   TTSLNLQKDE PNGRASDTAG QEQQYDVHGD VISAVVYQRF HVFGPEGKVF DGDAGGLTLP
121   GAGAFWGTLF TNDLQRLYKD TVSFQYNAVG PYLNINFFDS SGSFLGHIQS GGVSTVVGVG
181   GGSGSWHNA
```

Figure 6.16 Potential *O*-acylation sites of VapA

Each coloured block was considered separately in the design of primers for site-directed mutagenesis (as shown in table 6.3).

Table 6.3 Targets for Alanine-scanning mutagenesis of VapA *O*-acylation sites

Block ID (Mutant designation)	Amino acyl sequence changes	
	Native	Mutant
Block 2 (S36A)	VLD**S**GSSSAI	VLD**A**GSSSAI
Block 3 (38SSS-AAA40)	LDSG**SSS**AIL	LDSG**AAA**AIL
Block 4 (45SGAGS-AGAGA49)	LN**S**GAG**S**GIV	LN**A**GAG**A**GIV
Block 5 (54SGS-AAA56)	IVG**S**G**S**YDSS	IVG**A**G**A**YDSS
Block 6 (59SSTTS-A(5)63)	YD**SSTTS**LNL	YD**AAAAA**LNL
Block 7 (76SDT-ADA78)	NGRA**S**D**T**AGQ	NGRA**A**D**A**AGQ
Block 8 (S93A)	GDVI**S**AVVYQ	GDVI**A**AVVYQ
Block 9 (T118A)	AGGL**T**LPGAG	AGGL**A**LPGAG
Block 10 (128TLTF-ALFA131)	FWG**T**LF**T**NDL	FWG**A**LF**A**NDL
Block 11 (141TVS-AVA143)	YKD**T**V**S**FQYN	YKD**A**V**A**FQYN
Block 12 (160SSGS-AAGA163)	FFD**SS**G**S**FLG	FFD**AA**G**A**FLG
Block 13 (S170A)	GHIQ**S**GGVST	GHIQ**A**GGVST
Block 14 (174ST-AA175)	GGV**ST**VVGVG	GGV**AA**VVGVG
Block 15 (183SGS-AGA185)	GGG**S**G**S**WHNA*	GGG**A**G**A**WHNA*

In all cases, targets were successfullly mutated to encode alanines using mutagenic primers described in table 2.5. *Corynebacterium glutamicum* was transformed with the *pEKEx2-vapA* mutants and the subcellular localisation of the mutant protein was investigated in each case. The results revealed no significant change in the distribution of VapA (Appendix H). Logical analysis of this outcome did not exclude a covalent lipid/mycolic acid modification of VapA. That no changes in

compartmentalisation were seen in any of the mutants might indicate that a mycolic acid modification is not absolutely critical with regard to the retention of VapA in wall material within this heterologous system. An alternate explanantion is that a unique mycolylation is sufficient for retention but does not represent the biological norm, *i.e.* multiple sites are lipidated. However, with the significant circumstantial evidence suggesting a co-purification of protein–bound mycolates and VapA in *C. glutamicum* and *R. equi*, an *in vitro* assay was sought that might provide a superior analysis.

6.7 Analysis of mycoloyl transferase activity in VapA acylation

Recently, a study of porin mycoloylation activity across a panel of *C. glutamicum* mutants demonstrated that mutation of *Cg0413*, which encodes the mycolyltransferase Cmt1, was the cause of a loss-of-function in this respect (Rath *et al.*, 2013). Two approaches were taken with the aim of confirming the circumstantial evident for VapA mycoloylation *via* an *in vitro* enzyme assay. Firstly, it was considered that the activity of Cmt1 might be explored with VapA as a terminal acceptor of mycolic acids derived from extractable glycolipid fractions of representative mycolata. Secondly, a bioinformatics approach could be taken to seek a Cmt1 orthologue(s) in *R. equi*.

6.7.1 Extraction of Corybacterium glutamicum genomic DNA

In order to amplify *Cg0413*, a high quality DNA preparation was required to serve as a template for polymerase chain reaction. The extraction of genomic DNA from *C. glutamicum* ATCC 13032 was performed as described in section 2.7.1. The result was DNA of high quality (> 10,000 bp) which migrated as a single band on agarose

gel electrophoresis (Figure 6.17). The DNA was not fragmented and there was no low molecular weight contaminating nucleic acid. The DNA was considered appropriate to be used as a template for PCR.

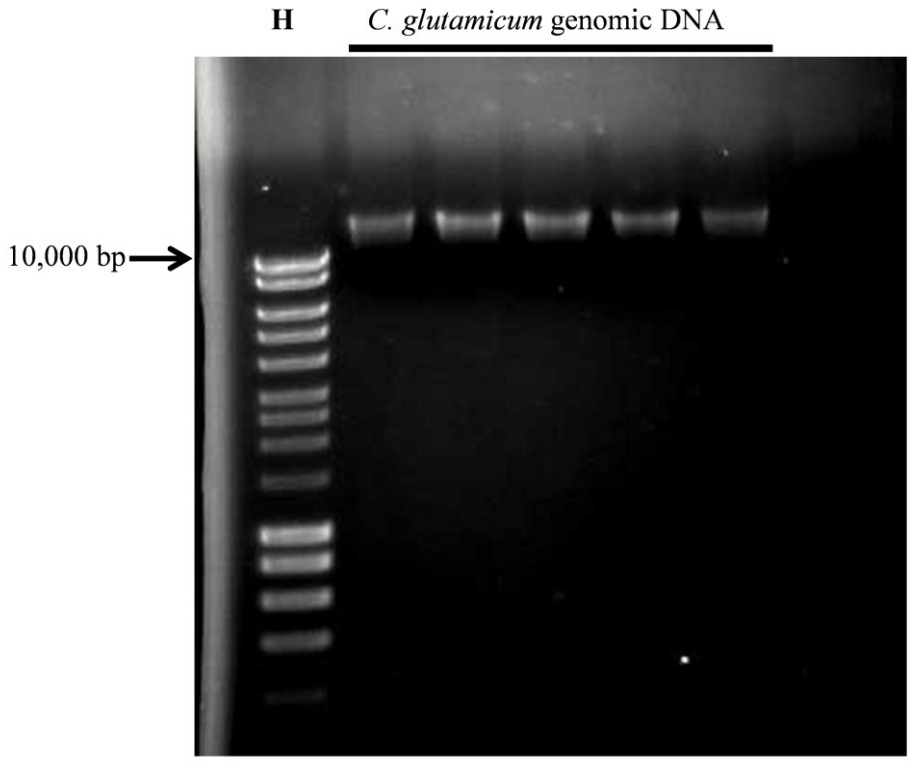


Figure 6.17 Analysis of *Corynebacterium glutamicum* genomic DNA
H represents hyperladder I.

6.7.2 Amplification of Cg0413

The open reading frame encoding corymycoloyltransferase 1 (designated *Cg0413*) was amplified without the signal peptide region for cloning into pET23a and pET28a for production with a *C*-terminal or *N*-terminal hexahistidine tag respectively. Electrophoresis revealed that the reaction generated a DNA band migrating consistently with the expected 1008 bp product (Figure 6.18). The product was purified from agarose and subjected to restriction digestion for cloning into vectors.

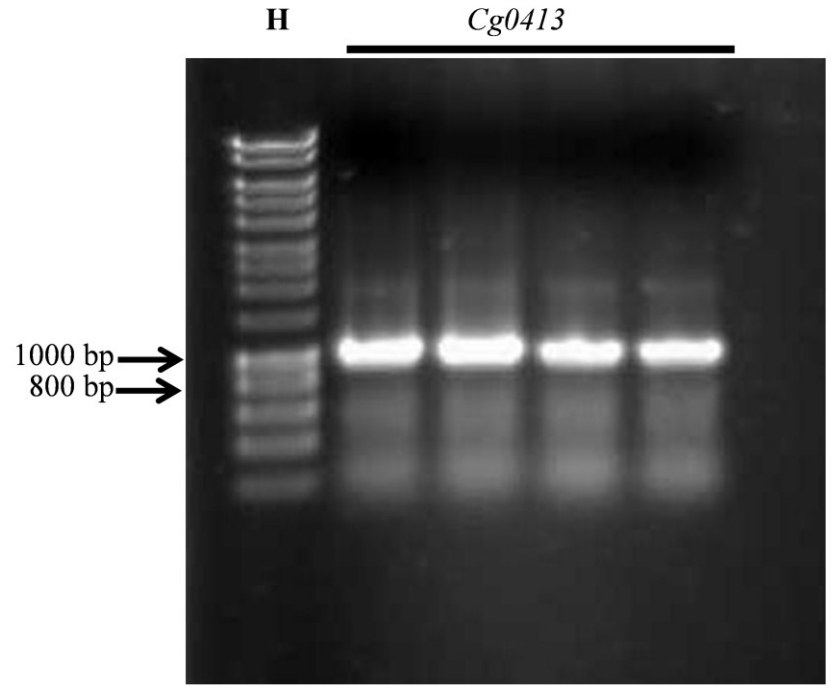


Figure 6.18 PCR amplification of *Cg0413*
PCR reaction was performed as described in sections 2.7.3 with oligonucleotide primers in table 2.1. H represents Hyperladder I.

6.7.3 Cloning of Cg0413 into vectors

The amplified *Cg0413* was purified, digested with the restriction endonucleases *Not*I and *Nde*I and purified from agarose. Thereafter the digested products were ligated to pET23a and pET28a (also digested with *Not*I and *Nde*I) as described in the methods section. The ligation mixture was used to transform competent *E. coli* Top 10. Plasmid DNA was recovered from antibiotic resistant transformants and screened for the release of *Cg0413* (migrating consistent with ~1000 bp fragment) after digestion with *Nde*I and *Not*I and subsequent electrophoresis (Figure 6.19). The sequence of the selected clones was confirmed by nucleotide sequencing.

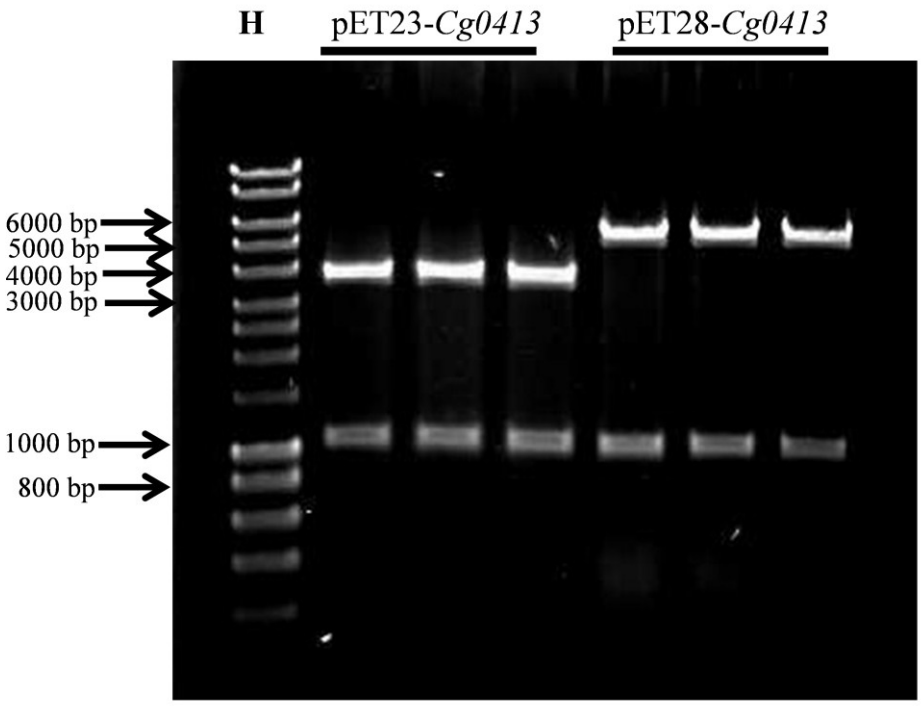


Figure 6.19 Restriction digestion analysis of recombinant plasmids pET23-*Cg0413* and pET28-*Cg0413*
The bands clearly reveal the successful cloning of the cmt gene into the plasmids.

6.7.4 Expression and purification of Cmt1 corynomycoloyl transferase

Sequencing results were in the affirmative since the gene was in frame with no mutation thus recombinant plasmid was used to transform *E. coli* BL 21 (DE3) cells. A single colony was grown in 10 mL LB broth (with antibiotic) as a starter culture overnight at 37°C with shaking at 200 rpm. This was used to inoculate 1 litre of terrific broth (with appropriate antibiotic) and grown at 37°C to OD_{600} of about 0.8. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM and incubated overnight at 20°C. After overnight incubation, cells were

harvested, re-suspended in lysis buffer, sonicated and clarified as described in section 2.8.5. Protein was purified in a single step *via* immobilised metal affinity chromatography as described in section 2.8.7.1. For both variants, there was significant gene expression. SDS-PAGE profile of peak fractions after IMAC for N-terminal hexa-histidine tagged protein is shown in figure 6.20.

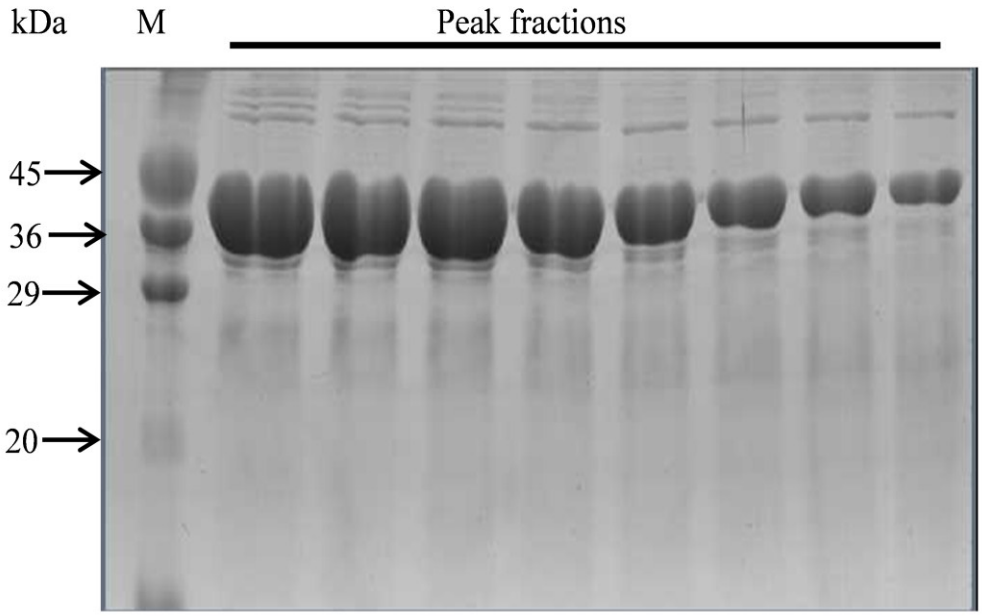


Figure 6.20 Peak fractions of Cmt1 subjected to SDS-PAGE following IMAC
'M' size markers; 20 kDa (Trypsin Inhibitor from soybean), 29 kDa (carbonic anhydrase from bovine erythrocytes), 36 kDa (glyceraldehyde-3-phosphate dehydrogenase from rabbit muscle), 45 kDa (ovalbumin from chicken egg).

6.7.5 Extraction of cell wall associated lipids

Extractable cell wall associated lipids were considered as suitable substrates for studying protein mycoloylation. Lipids that are non-covalently associated with the rhodococcal cell envelope include TDMs and TMMs. TMM has been used as a lipid donor in mycoloylation reactions catalysed (addition of mycoloyl residue from TMM

to another TMM molecule in the synthesis of TDM) by the *Mycobacterium tuberculosis* Ag85A which is an homologue of Cmt1 (Kremer *et al*., 2002; Elamin *et al*., 2009; Huc *et al*., 2013). Thus the surface lipids of both *Rhodococcus equi* and *Corynebacterium glutamicum* were considered as potential sources of mycolates for the lipidation of VapA in an *in vitro* model. Hence crude preparations of the surface attached lipids were made from *C. glutamicum* and *R. equi*. The composition of the extracts from each microorganism varied markedly. The extract from *R. equi* showed the majority to be in the form of TDM, while in the extract of *C. glutamicum*, TMM dominated (Figure 6.21).

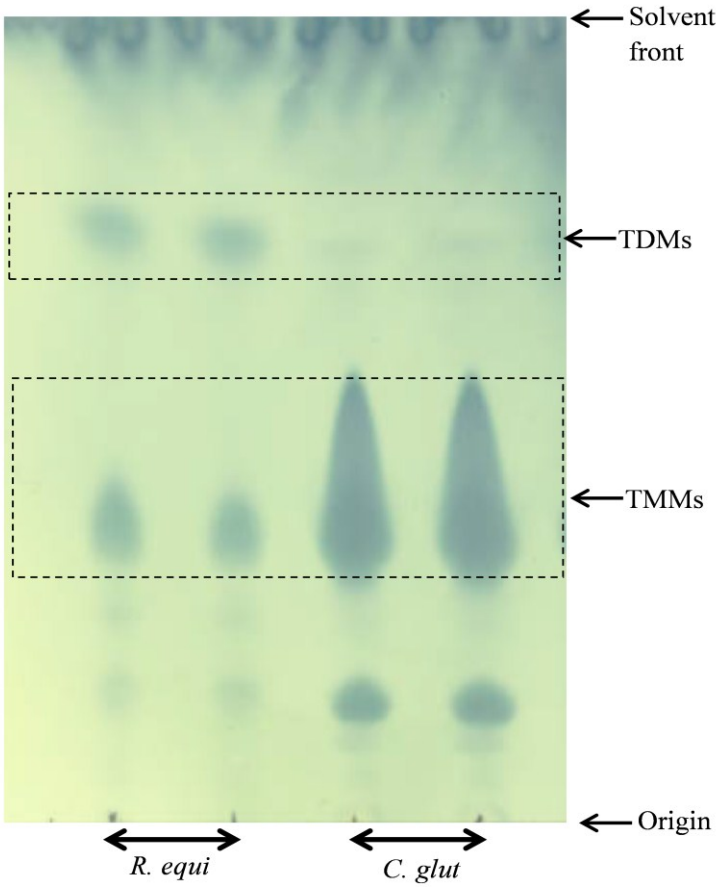


Figure 6.21 Compositonal analysis of envelope glycolipids of *R. equi* and *C. glutamicum*

Lipids extracted with chloroform:methanol and loaded on silica. TLC was developed with chloroform:methanol:water (60:16:2). Lipids were visualised by staining with 5 % ethanolic phosphomolybdic acid and charred. TDM (trehalose dimycolate) and TMM (trehalose monomycolate) were identified by comparison to similar profiles (Brand *et al.*, 2003)

6.7.6 Mycoloyltransferase assay

Cell wall associated lipids (dried) were re-suspended in buffer as detailed in section 2.10.1. In the complete reaction, the lipid suspension was mixed with VapA (1 mg/mL) and the reaction was initiated by the addition of Cmt1 (1 mg/mL) and incubated for 45 min as described in section 2.10.1. Enzymes tagged at both termini

were used with identical outcomes. Additionally, both variants of hexahistidine tagged mature VapA were used, again with identical outcomes. A series of controls excluding individual components were conducted simultaneously. Aliquots of the various reactions were subject to Urea PAGE which has been used to demonstrate various acyl:protein adducts (Brown *et al*., 2005).

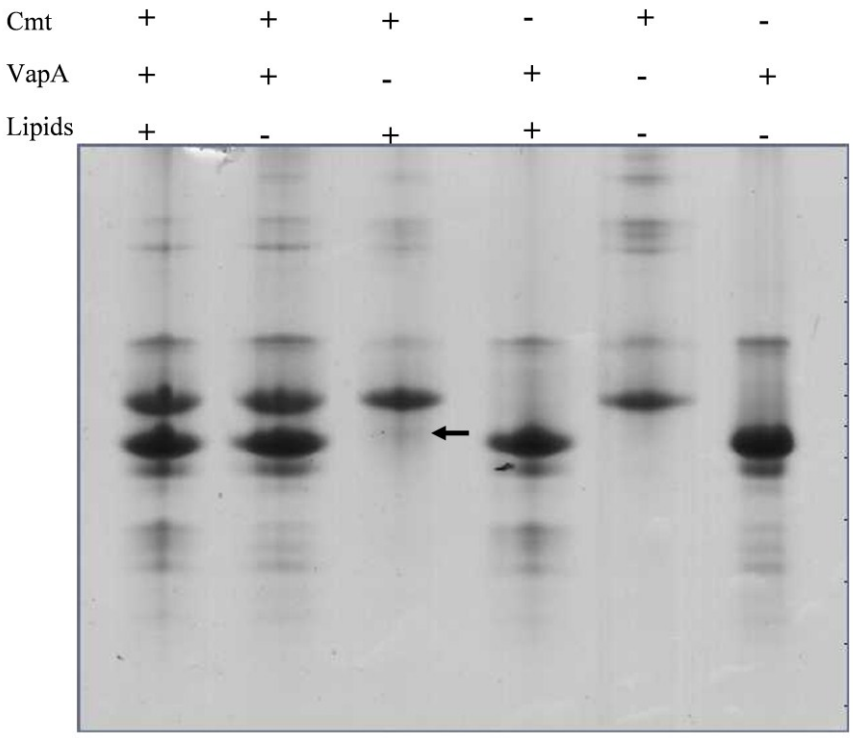


Figure 6.22 Urea-PAGE profile of Cmt assay using lipids from *C. glutamicum* as donors and VapA as an acceptor substrate
Volumes are 10 µL of Cmt (1 mg/mL) and VapA (1 mg/mL), 5 µL of lipid solution.

When the mycolyltransferase was incubated with VapA and lipids from *C. glutamicum,* no shift in the mobility of VapA or Cmt1 was evident (Figure 6.22). However comparison of the banding profiles generated by Cmt1 alone and the protein incubated with lipids revealed a faint but reproducible band (arrowed) migrating more quickly than the recombinant enzyme. This banding pattern could indicate the limited formation of an acyl enzyme intermediate at an active site serine residue consistent with the recognised mode of action of mycobacterial mycolyltransferases (Belisle *et al*., 1997) and other esterases. Its presence in the complete assay would likely have been masked by the VapA component.

When similar reactions were carried out using *R. equi* lipids, a similar shift in migration was evident but this putative acyl enzyme intermediate (arrowed) appeared far more abundant in this instance when using heterologous lipids (figure 6.23), although there is a recognised overlap in acyl chain length (Sydor *et al*., 2013). Again no shifts in the migration of VapA were noted.

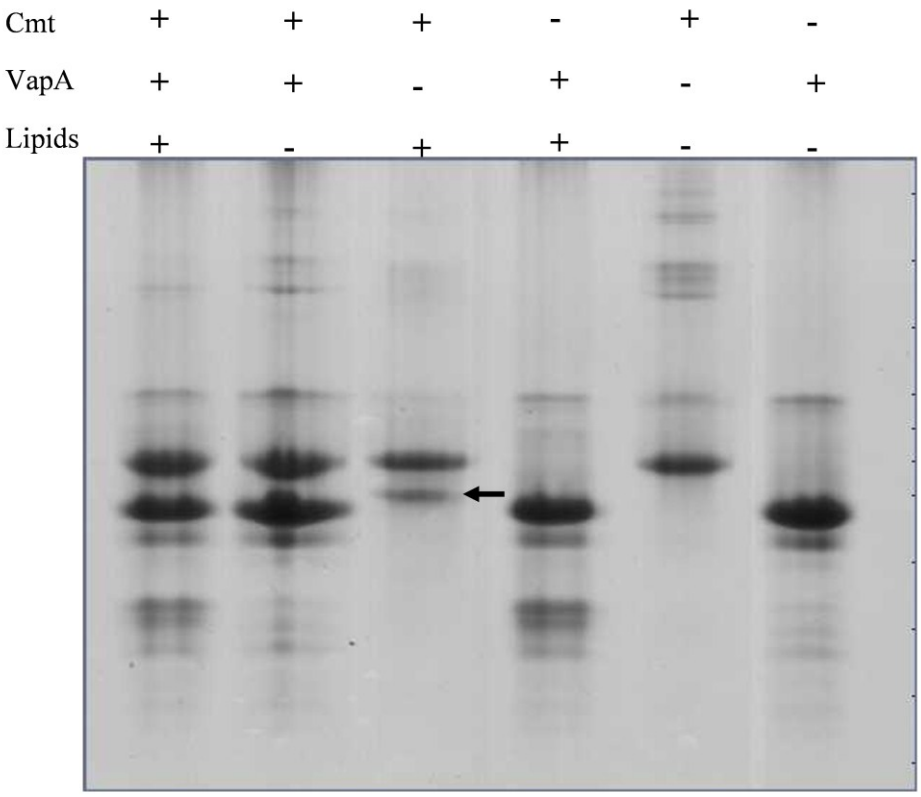


Figure 6.23 Urea-PAGE profile of Cmt assay using lipids from *R. equi* and VapA as substrates
Volumes are 10 µL of Cmt (1 mg/mL) and VapA (1 mg/mL), 5 µL of lipid solution

Encouraged by the possibility of the formation of an acyl:enzyme intermediate, efforts were made to refine this assay approach. The enzyme was incubated with the lipids from both sources in different proportions and was analysed by Urea-PAGE as

before. It was hoped that optimising lipid loading to the enzyme might enable detection of lipid transfer to VapA in the final stage of the complete reaction.

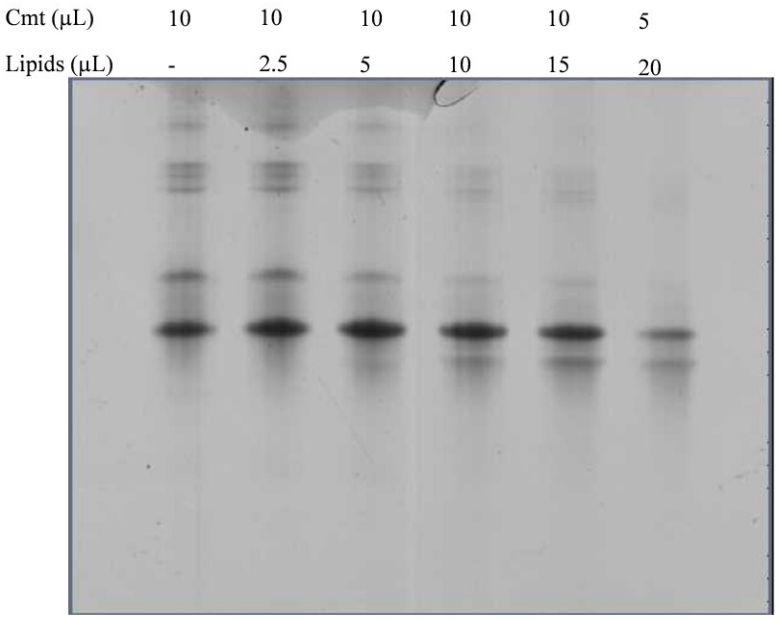


Figure 6.24 Optimisation of adduct formation between Cmt and cell wall associated lipids from *C. glutamicum*.
Crude lipids were solubilised in Tris-DTT and incubated with Cmt (1 mg/mL) as indicated.

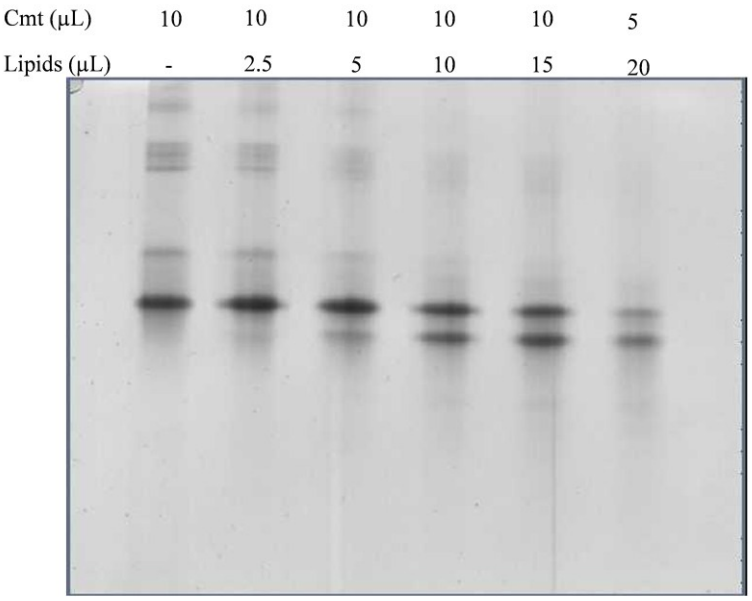


Figure 6.25 Urea-PAGE profile of optimisation of adduct formation between Cmt and cell wall associated lipids from *R. equi*.
Crude lipids were solubilised in Tris-DTT and incubated with Cmt (1 mg/mL) as indicated

Figures 6.24 and 6.25 show that the lipids from both systems appeared to form apparent adducts with the enzyme. This was dependent on the volume of lipid suspension applied and thus appeared concentration dependent. Reducing the volume of the enzyme (5 μL) also reflected on the adduct formation. In both cases, optimum conditions were found when 10 microlitres of protein (1 mg/mL) were mixed with 15 microlitres of lipids. It is interesting to note that the banding profile of Cmt1 on Urea-PAGE is complex, especially in the absence of lipid. Here multiple slowly migrating forms are evident relative to the modal state. Titration with lipids refined the profile with the less abundant forms appearing to unify into the putative adduct band with a relatively smaller impact on the abundant form of the enzyme. Attempts to extract lipids from the recovered adduct as methyl esters (as with all other extractions of potentially acylated proteins from polyacrylamide in this study)

yielded no detectable MAME on MALDI-TOF analysis. This could be interpreted as indicating a lack of lipid in this material, its induced novel migration being the result of a lipid-driven denaturation of the enzyme, or simple poor recovery of *bona fide* acyl adducts from gels.

The complete mycoloyltransferase assays were carried out with optimised enzyme:lipid composition being combined with VapA as a terminal acceptor for the lipids. In the first instance, all the components *i.e.* VapA, lipids, and Cmt were added at the same time while in the other case, Cmt and the lipids were incubated for 30 min before addition of VapA and incubation for a further 30 min. Analysis on Urea-PAGE revealed that no modification to the migration of VapA was seen with either methodology (Figures 6.26 and 6.27).

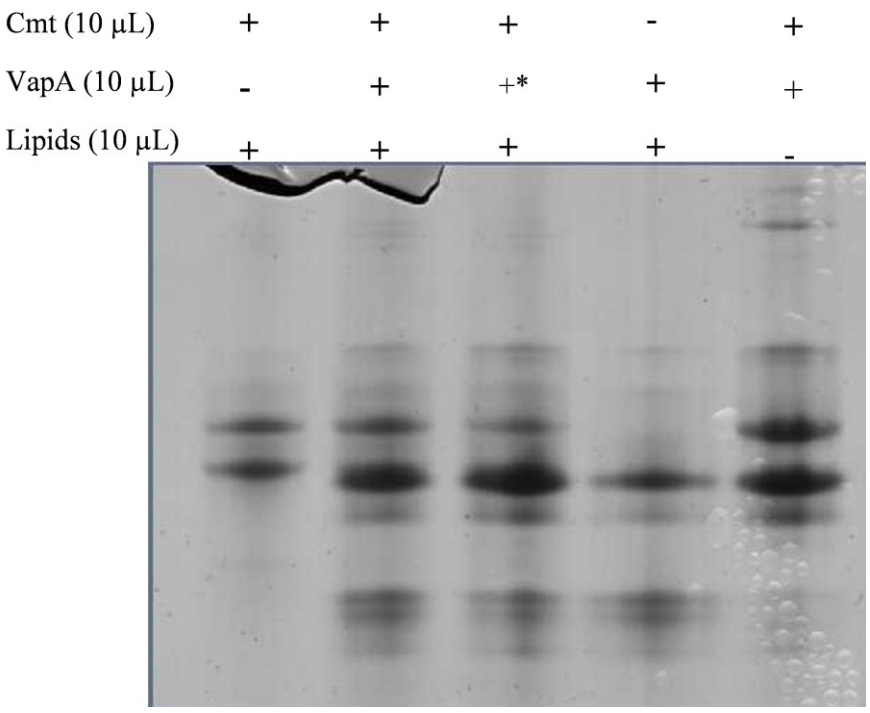


Figure 6.26 Urea-PAGE profile of Cmt assay using lipids from *C. glutamicum* and VapA as substrates

Volumes are 10 µL of Cmt (1 mg/mL) and VapA (1 mg/mL), 10 µL of lipid solution. The asterisk indicates that Cmt and lipids were incubated for 30 min before adding VapA

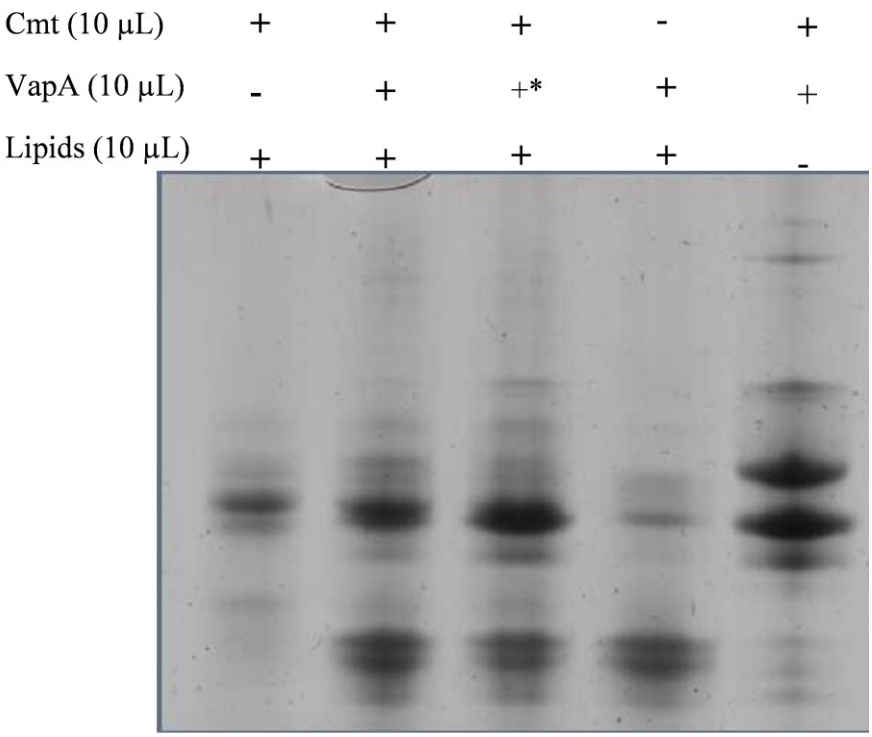


Figure 6.27 Urea-PAGE profile of Cmt assay using lipids from *R. equi* and VapA as substrates
Volumes are 10 µL of Cmt (1 mg/mL) and VapA (1 mg/mL), 10 µL of lipid solution. The asterisk indicates that Cmt and lipids were incubated for 30 min before adding VapA

Several possible explanations for the apparent lack of transfer of lipids to VapA exist. Firstly, the hypothesis that Cmt1 can mycolyolate VapA *in vivo* might be rejected, the body of circumstantial evidence being rejected as artefact. However, explanations that conserve the hypothesis remain. It is possible that (i) the recombinant Cmt1 preparation is inactive and the various putative mycoloyl adducts are artefactual, (ii) the Cmt1 preparation is active and the various putative mycoloyl adducts are real but VapA is presented in a non-accepting configuration and (iii) the composition of the urea gel is appropriate for revealing formation of an acyl:Cmt1 intermediate but not for the resolution of a VapA adduct.

In order to explore these possibilities *via* a more sensitive and discriminating assay, ^{14}C radiolabelled lipids from *Mycobacterium bovis* BCG (a gift from Prof G. Besra) were used in the similar mycoloylation assays.

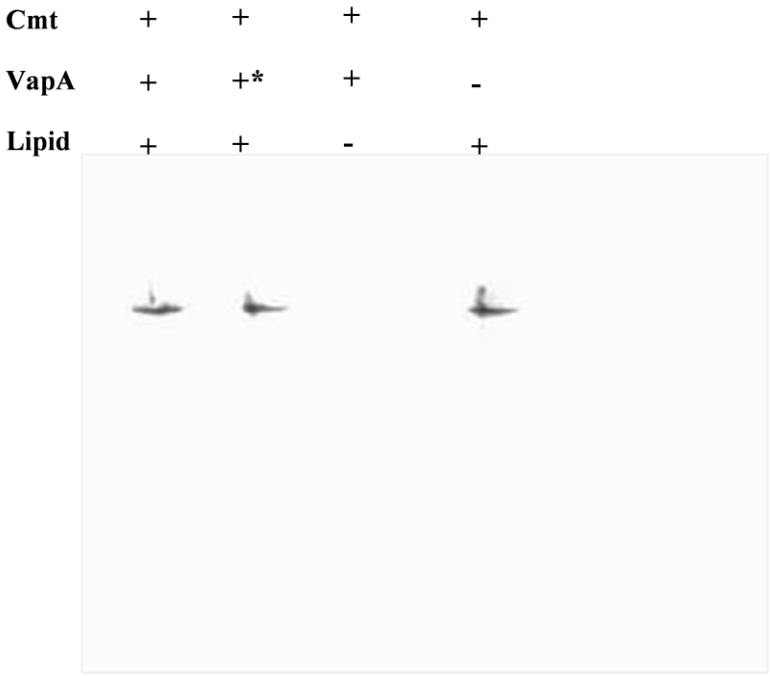


Figure 6.28 Autoradiogram Urea-PAGE profile of Cmt assay using radiolabelled lipids from Mycobacterium bovis BCG and VapA as substrates

Volumes are 10 µL of Cmt (1 mg/mL) and VapA (1 mg/mL), 10 µL of lipid solution. The asterisk indicates that Cmt and lipids were incubated for 30 min before adding VapA

Autoradiography (Figure 6.28) of a Urea-PAGE gel indicated the presence of a single radio-labelled species with a similar rate of migration to the apparent adduct seen with lipids from other sources. This is interpretted as indicating a significant stable association of *M. bovis* lipids with the mycolyltransferase and essentially

rejects (i) above. This would also indicate that Cmt1 might have a very relaxed specificity for mycoloyl substrate chain-length. However, as this experiment did not generate any other bands associated with expected VapA migration, it revealed no evidence for any lipid transfer to VapA, rejecting (iii) above.

It remains possible that the conformation of the accepting protein is critical for activity. When one considers that lipoprotein processing is simultaneous with plasma membrane secretion (Sutcliffe and Harrington, 2002) then a folded purified (and artificially tagged) protein might not readily interact with an active Cmt1 preparation. It is notable that an *in vitro* demonstration of Cmt1 modification of its only experimentally-validated substrate (PorA) was not a feature of its recent characterisation (Rath *et al.*, 2013). Exploration of small peptide fragments of VapA and PorA as acceptor substrates that might emulate less-structured forms of those proteins would be warranted.

6.7.7 Identification of putative mycoloyltransferase of Rhodococcus equi 103S

During the conduct of this study the product of *C. glutamicum cmt1* (Cg0413), an already recognised mycolyltransferase (Brand *et al*., 2003) was demonstrated to carry a protein mycolyltranferase activity (Huc *et al*., 2013). Furthermore, orthologues were identified from *Mycobacterium* and *Rhodococcus erythropolis* RP4 through their ability to complement an in-frame *cmt1* deletion mutation evidenced by the restoration of mycolylation of the porin substrate proteins CgPorAB. In order to establish the identity of any *Rhodococcus equi* orthologue and study its mycolyltransferase activity with respect to VapA modification, Cmt1 was used as a query in BLASTP (Altschul *et al*., 1990) searching the *R. equi* 103S proteome. Four

sequences were recognised with alignment scores =>200 and E values < $4x10^{-63}$ (Table 6.4).

It is likely that the mycolylation of proteins is a critical component process that contributes to the population of the mycolate membrane with proteins. The molecular basis of protein trafficking and insertion is inherently interesting given the unusual composition of these membranes. Study of the three genera *Rhodococcus, Mycobacterium* and *Corynebacterium* offers opportunities for comparisons of physical properties and recognition of any attendant biological modifications that may greatly extend our current understanding of the function of this key biological structure. All three genera include significant pathogens. It is well recognised that surface structures are critical to key aspects of infection including adherence, nutrient acquisition, cellular and tissue invasion, and immune modulation. Experimental definition of substrate protein recognition motifs might help establish a surface-exposed target for anti-infective discovery. In order to achieve this, a biochemical characterisation of protein mycolylation is required. In order to facilitate this study in *R. equi*, a strategy was required to identify candidate genes for cloning and product characterisation.

A feature of all of the mycolata is their apparent redundancy in mycolyltransferase activity. Before its association with protein acceptors of mycolic acids, Cmt1 was shown to participate in the synthesis of the trehalose dicorynomycolate glycoconjugate. Given this apparent redundancy and the promiscuity exhibited by individual enzymes in terms of acceptor specificity, it is important to carefully consider sequence data before defining characterisation studies. In order to establish strong candidacy for orthology, in terms of its protein mycolyltransferase activity, it

is important to establish that homologues generated in this initial BLASTP screen identify Cmt1 as best match when used as a query. The protein sequence for each of the *R. equi* 103S 'hits' was then used as a query in a BLASTP search of the *C. glutamicum* ATCC 13032 proteome. Interestingly, these reciprocal analyses revealed that each of the 4 candidate *R. equi* mycolyltransferase sequences shared strongest homology with *C. glutamicum* Cmt1 (NP_599594.1), which given the most trivial interpretation, might suggest significant redundancy in provision for protein mycolyltransferase activity in *R. equi*.

The amino acid sequence of *R. erythropolis* PR4 RER_15370 (Genbank PID YP_002764984.1), which restored protein mycolyltransferase activity in the *C. glutamicum* Δ*cmt1* mutant, was used as a query in a BLASTP search limited to *R. equi* 103S. This search generated nine hits with alignment scores =>200 and e scores $< 5x10^{-77}$ (Table 6.5).

The existing annotations associated with these hits were of two types: esterase or secreted esterase. It is well established that mycolyltransferase activity is associated with the pseudoperiplasm and this had been recognised for Cmt1 through its possession of a signal peptide leader sequence. Two of these interesting hits, namely REQ_32810 and REQ_02570 (ranked 1 and 4 respectively, Table 2) were annotated as esterases. Individual scrutiny of both of these archived predicted protein sequences using SignalP4.1 (Peterson *et al*., 2011) failed to identify a canonical signal peptide for REQ_32810 but did recognise REQ_02570 as a putative secreted protein. It is noteworthy that SignalP3.0, which was current at the time of annotation predicted a non-secreted protein.

Table 6.4 Identification of *Rhodococcus equi* protein mycolyltransferase candidates based on Cmt1 homology.

R. equi 103S BLASTP hit		E value	Reciprocal *C. glutamicum* BLASTP hit		E value
1	YP_004007959.1 REQ_32180	$1x10^{-85}$	1	NP_599594.1	$4 x10^{-87}$
			2	YP_225214.1	$6 x10^{-64}$
			3	NP_600254.1	$3 x10^{-49}$
			4	YP_227116.1	$2 x10^{-48}$
2	YP_004008131.1 REQ_34590	$7 x10^{-78}$	1	NP_599594.1	$6 x10^{-83}$
			2	YP_225214.1	$6 x10^{-61}$
			3	YP_227116.1	$2 x10^{-48}$
			4	NP_600254.1	$4 x10^{-47}$
3	YP_004005782.1 REQ_09940	$9 x10^{-70}$	1	NP_599594.1	$8 x10^{-70}$
			2	YP_225214.1	$2 x10^{-54}$
			3	NP_600254.1	$3 x10^{-45}$
			4	NP_602067.1	$8 x10^{-42}$
4	YP_004005093.1 REQ_02570	$7 x10^{-67}$	1	NP_599594.1	$7 x10^{-67}$
			2	YP_227116.1	$3 x10^{-60}$
			3	NP_600254.1	$2 x10^{-57}$
			4	NP_601385.1	$9 x10^{-42}$

The amino acid sequence of the experimentally-validated protein mycolyltransferase *C. glutamicum* Cmt1 was used as a BLASTP search limited to *R. equi* 103S. Candidates with strong alignment scores (+>200) (identified by Genbank PID) were used as BLASTP queries of *C. glutamicum* ATCC 13032.

Table 6.5 Identification of *Rhodococcus equi* protein mycolyltransferase candidates based on RER_15370 homology.

Rank	Description	Total score	Query cover	E value	Identity	Accession
1	esterase	486	89%	$8x10^{-172}$	72%	YP_004007959.1 REQ_32810
2	secreted esterase	401	81%	$1x10^{-139}$	69%	YP_004008131.1 REQ_34590
3	secreted esterase	382	99%	$1x10^{-131}$	55%	YP_004005782.1 REQ_09940
4	esterase	286	96%	$2x10^{-94}$	44%	YP_004005093.1 REQ_02570
5	secreted esterase	285	97%	$4x10-^{94}$	46%	YP_004008589.1 REQ_39250
6	secreted esterase	280	84%	$3x10^{-92}$	48%	YP_004008523.1 REQ_38570
7	secreted esterase	249	80%	$7x10^{-80}$	44%	YP_004006071.1 REQ_12940
8	secreted esterase	248	86%	$4x10^{-79}$	41%	YP_004009020.1 REQ_43790
9	secreted esterase	242	85%	$5x10^{-77}$	42%	YP_004006331.1 REQ_15690

The amino acid sequence of the experimentally-validated protein mycolyltransferase *Rhodococcus erythropolis* PR4 RER_15370 was used as a BLASTP search limited to *R. equi* 103S. Statistics describing candidates with strong alignment scores (+>200) are tabulated.

However, analysis of sequence alignments (AlignX) revealed that REQ_32810 was predicted to carry a significant *N*-terminal extension of between 47 to 60 amino acids relative to Cmt1 and *R. equi* homologues (Figure 6.29). Five other potential start

codons (A/G-TG→V/M) could be considered for REQ_32810, two coincide with positions in the alignment that correspond to predicted *N*-termini of REQ_02570 and Cmt1 respectively, while one coincides with that for both REQ_09940 and REQ_34590.

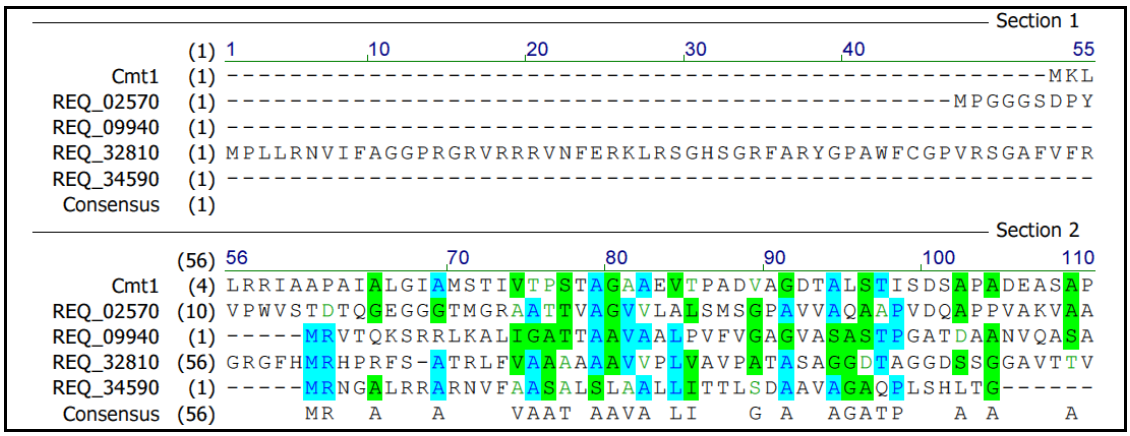


Figure 6.29 Alignment of predicted gene products representing closest *R. equi* homologues of *C. glutamicum* Cmt1 used to explore alternate *N*-terminus for REQ_32810.

Alignments were generated using the AlignX within Vector NTI (Invitrogen)

Analysis of predicted REQ_32810 gene products initiated at each of the alternate initiation codons using SignalP4.1 revealed a likely signal peptide if the *N*-terminus occupied position 61 by the current numbering, which would be coincident with the *N*-termini of both REQ_09940 and REQ_34590. On the basis of a likely misinterpretation of the site of translation initiation and its strongest homology with Cmt1 of *R. equi* gene products, REQ_32810 was retained in the set of candidate protein mycoloyltransferases; in subsequent sequence alignments, residue M61 was assumed the *N*-terminus of the nascent REQ_32810 protein.

When the amino acid sequences of each of the 9 RER_15370 homologues were used as a BLASTP query of *C. glutamicum* ATCC 13032, those ranked 1-6 (Table 6.5) all showed greatest homology with Cmt1. The difference in the extent of amino acid identity between those ranked 2 (69%) and 3 (55%) is marked and might indicate some resolution in biochemical role and substrate selectivity. Those ranked 1 and 2, REQ_32810 and REQ_34590 respectively share 72% and 69% sequence identity with respectively RER_15370 and 66% with each other. Both might be considered strong candidates, efforts to clone these enzymes toward the end of this study were unsuccessful.

CHAPTER SEVEN

7 General discussion

Despite research into *Rhodococcus equi* pathogenesis, rhodococcosis remains a major cause of wastage in the equine industry. Though foals are able to develop protective immunity against infection, there are still no registered vaccines against rhodococcosis (Lohmann *et al*., 2013). This lack of vaccines could be primarily attributed to the complexity of the immunity to *R. equi*; the disease affects foals under six months of age, thus a vaccine must provide protection against *R. equi* infection early in life (Bordin *et al*., 2013).

Because of the intracellular nature of the organism, the range of antibiotics that could be effective in the treatment of *R. equi* infections is limited. Treatments involving combinations of rifampin (or doxycycline) with other antibiotics such as azithromycin, clarithromycin has been recommended but have produced mixed results (Giguere *et al*., 2012; Venner *et al*., 2013). However, the emergence of antibiotic-resistant strains which tolerate macrolides and rifampin has led to concern around the definition of an effective treatment strategy (Yamshchikov *et al*., 2010; Burton *et al*., 2013).

The ability of *Rhodococcus equi* to survive in the harsh macrophage environment is, in part, defined by its complex cell envelope. The wall's configuration offers some resistance to antibiotics. Mycolic acids present a formidable permeability barrier and channel-forming porins control access through their selectivity. The biosynthesis and assembly of its dominant components such as mycolic acids and arabinogalactan could represent useful drug targets (Prescott, 1991; Sutcliffe *et al*., 2010).

More broadly, both chromosomal and plasmid-encoded factors have been identified as potential drug targets for *Rhodococcus equi* infections. Important chromosomal-encoded factors include the enzymes of aromatic acid biosynthesis (Letek *et al*. 2010,), enzymes of steroid catabolism (van der Geize *et al*., 2011), nitrate reductase (Letek *et al*., 2010) and rhequichelin (Miranda-CasoLuengo *et al*., 2012). The plasmid encoded factors are mainly proteins (which are secretory in nature) expressed by genes of the *vap* cluster of the pathogenicity island (Meijer and Prescott, 2004).

Secretory proteins especially of intracellular pathogens are often in direct contact with host tissues and thus are of utmost importance in the investigation of bacterial pathogenesis. *Rhodococcus equi* expresses a group of Virulence associated proteins (Vaps) that are encoded by the *vap* locus of the plasmid-borne pathogenicity island which contributes significantly to pathogenesis. Almost all Vaps are secreted, diffusible and immunogenic and thus are primary targets for vaccine development, but VapA (a surface-located protein) been shown to induce the strongest immune response in both foals and adult horses (Kaufmann and Hess, 1999; Byrne *et al*., 2008; Dawson *et al*., 2010).

VapA has been shown to be a major virulence determinant since mutants that lack its gene were attenuated to the same degree as a mutant that lacks the majority of the *vap* locus. Nevertheless VapA alone is not sufficient to support the ability to replicate in the macrophage (Jain *et al*. 2003, Vazquez-Boland *et al.*, 2013); that may require interplay of various factors. Despite many years of study, the mechanism by which Vaps exert their effects remains unclear. One avenue for research was to

address the lack of understanding of the structural basis for virulence; this was the main focus of chapter four.

The apparently dominant role of VapA in virulence, despite its limited mobility in comparison to other closely-related Vaps, remains interesting. The tethering of this protein at the cellular surface was considered as an interesting rationale for its ascendency. Disruption of this tethering may limit virulence and extend knowledge regarding its mode of action. In this regard, identifying the nature and the site of the lipid modification evidenced elsewhere by radiolabelling with ^{3}H-palmitate (Tan *et al*, 1995) was a major goal (Chapter six).

It was also considered that the Vaps might act in a concerted fashion, their similarity might promote assembly of a cell surface structure tethered by VapA to the mycolic acid layer of the outer wall. In order to explore this hypothesis, a biophysical approach to monitoring Vap interactions was taken in Chapter five.

7.1 Production, purification and structural characterisation of Vaps

Apart from VapC, all of the Vaps of *Rhodococcus equi* 103S were produced in *E. coli* and purified using a combination of metal chelate affinity chromatography followed by anion exchange chromatography. Given the overall similarity of this family of proteins, the lack of success in achieving reasonable yields of VapC is puzzling. In order to make progress, VapC was omitted from further studies but it would be valuable to explore the reasons for the poor yield in future. Modifications to replace any uncommon *E. coli* codons or to remove secondary structures in mRNA that might inhibit translation (Usha *et al*., 2006) could be considered.

Although all of the *C*-terminally hexahistidine-tagged Vap proteins tested concentrated well, only VapG crystallised. Several crystal forms were achieved but one diffracted well to 1.8 Å. At this time we became aware that VapD had also been crystallised, its structure was resolved first (Whittingham *et al*., 2014) via the addition of extra methionine residues wherever those were found in other Vap proteins. This led to resolution of phase *via* analysis of selenomethionine containing crystal. This success accelerated solving of VapG by molecular replacement based on the new VapD structure. Both structures revealed a core structure of an 8-stranded closed β-barrel representing the *C*-terminal part of the molecule. The *N*-terminal sequences in both cases appeared to be missing but given that the crystallisation of intact pure protein was rapid, they appear likely to be unstructured, rather than having been lost *via* degradative processes in each case. While solving the crystal structure, we also became aware that crystal structure of VapB had also been solved. All three proteins possessed a similar β-barrel topology, which has not been found in any other protein from other organisms. Although many outer membrane proteins from Gram-negative bacteria are β-barrel shaped (Mizianty and Kurgan, 2011), these barrels are typically embedded in the bacterial outer membrane having a hydrophobic exterior face and a more polar internal aspect and represent water-filled channels (Rigel and Silhavy, 2012). The uniqueness of the Vap structure could be an important feature for this group of proteins.

VapG differed from VapD in that two molecules were present in the asymmetric unit of the crystal. The barrel of that designated as molecule 1 exhibited a bound metal ion, which appears to be potassium, the corresponding site of molecule 2 was unoccupied. Whether the potassium ion is of physiological relevance is not known

since the mother liquor contained potassium chloride. Most of the co-coordinating groups were related to main chain atoms and a water molecule but mutation of residue Asn 107 and Asn 114, which had side chain ligands, combined with biochemical techniques such as isothermal titration microcalorimtery would allow investigation of specificity and affinity.

Homology modelling of the other Vap proteins revealed that they conformed well to this topology. Although no functional insight can be immediately drawn from these structures, the bipartite organisation into an unstructured *N*-terminal region and a highly conserved closed β-barrel core appears to be a common feature of these proteins. The flexibility that this *N*-terminus exhibits might allow it to interact freely with other proteins, either host-derived or bacterial, including other Vaps. It seems less likely to represent an anchoring device to the mycolate membrane as it is not restricted to VapA and VapB. Now that the boundaries of these domains are understood, these can be dissected through molecular approaches to define individual functions and synergies using both infection based and biochemical investigations. In terms of the structured core domain, we now possess a valuable blue-print for structure: function studies.

Chapter five also reported the effect of metals on the hydrodynamic behaviour of VapA and VapG using analytical ultracentrifugation. While both proteins (at 1 mg/mL) exist in a monomer-dimer equilibrium, K^+, Ca^+, Li^+ seem to reduce the tendency of VapG to dimerise. In both cases, the apparent molecular masses were found to be higher than the calculated molecular masses of each recombinant molecule which could be ascribed to the possession of a disordered *N*-terminus. The self-association of VapA might be attributed partly to the intrinsic ability of the

protein to dimerise (since the protein exists as a monomer and a dimer) while the self-association of VapG was caused by the forces of sedimentation. This might have led to the effect of the metals on the hydrodynamic traits of the proteins.

7.2 Interaction of Vaps in free solution with immobilised VapA

Since VapA alone could not confer on the organism the ability to proliferate in macrophages, and expression *vapA* and other Vap genes is under the control of the *VirR* operon (Wang *et al*., 2014), functional interaction between VapA and other Vap proteins was possible. In chapter five, the study sought to determine whether Vaps in free solution could interact with an immobilised VapA. It was demonstrated *via* surface plasmon resonance that VapA could interact with all the other Vap proteins (except VapD). Vap G associated and dissociated rapidly while VapE and VapH associated and dissociated at a slower rate. The low micromolar equilibrium dissociation constant is consistent with a transient interaction. These observations are very preliminary; the study will be extended to address several considerations. The immobilisation of VapA could have occurred via two lysine residues, one in the unstructured *N*-terminus and the other in the β-barrel core. Alternative immobilisation strategies could utilise *N*- and *C*-terminal hexahistidine-tagged ligand *via* these tags on NTA-chips to explore the consequences of immobilisation mode. As VapA is induced at acidic culture pH, the effect of pH on these interactions should also be explored. Likewise, the impact of possible binding of K^+ by VapG could be explored by incorporating the metal in the buffers. Additionally, it is also possible that multi-Vap assemblages could form and that associations might be stabilised by subsequent binding of other partners. All of these scenarios can be explored using the SPR methodology.

7.3 Lipidation of VapA and other envelope proteins

Chapter six focussed on the lipidation of envelope proteins of *Rhodococcus equi*. Available evidence suggests that VapA is lipidated as it incorporates radiolabelled ^{3}H-palmitate (Tan *et al*., 1995), but it seems to be an unusual protein since its lipidation may be independent of lipoprotein diacylglycerol transferase (Lgt), which is a key enzyme in lipoprotein processing (Meijer and Prescott, 2004).

Some lipoproteins are non-covalently associated with the mycoloyl layer of the cell wall and, in the case of *C. glutamicum* porins, the proteins are covalently linked to mycolates. In the latter case, it has been reported that such mycoloylation is critical for those proteins to function (Huc *et al*., 2010). In fact many porins discovered in mycolic acid containing bacteria are integral membrane proteins thus could be lipidated (Siroy *et al*., 2008; Rath *et al*., 2011). Membrane proteins were extracted from *Rhodococcus equi* and VapA was confirmed to be among them. Further chemical analysis revealed that collectively, these proteins possess mycoloyl residues. When *vapA* was over expressed in *C. glutamicum* an increase in lipidation was seen. Although the study demonstrated protein mycolylation in *Rhodococcus*, a formal link to Vap was not achieved. Attempts to analyse lipids released from proteins recovered from polyacrylamide gels were unsuccessful.

Although the evidence suggested that mycolylation of VapA was a distinct possibility, a bioinformatic approach to identify a likely lipidation site was not possible. Although *C. glutamicum* PorA/H are lipidated *via* a mycolyl residue (Rath *et al*., 2013) they do not share any obvious sequence motif with the VapA/B. Initially Thr31 was identified as the putative lipidated residue in VapA as it occupied the position immediately after the predicted signal peptidase cleavage point

analogous to the cysteine residue usually modified in typical lipoproteins. As the link between lipid modification and membrane retention was anticipated to be critical, analysis of subcellular localisation was considered as a means to relate structure to function. Mutating Thr31 to Ala did not affect the subcellular localisation of the protein in *C. glutamicum*. All other threonines and serines were mutated and equally this did not change the subcellular localisation of VapA in *C. glutamicum*. In the light of structural data, the hydrophobic 'top' of the Vap molecules emerged as an alternate means of associating with the mycolate membrane. Alternatively, multiple mycolylation sites might be deployed and loss of one if the protein were in a dimeric form, might not equate to a loss of membrane association.

Recently a different strategy was used to investigate site of mycoloylation of PorH in *Corynebacterium glutamicum*. This involved a combination of 2D and 3D NMR through ^{1}H and ^{15}N protein assignment and partial ^{13}C assignment of protein residues and mycolic acid protons and carbons combined with MALDI-TOF MS (Rath *et al*., 2013). A similar strategy could be adopted to investigate the site for mycoloylation of VapA. Though lipidations on threonine and serine have been widely reported for membrane anchored lipoproteins, tyrosine is also a possible residue (Rath *et al.*, 2013).

In the structured barrel core of the VapA model, the side chains of Tyr 66, Thr 97, Ser 132, Ser 152 and Ser 154 (which occur in sheet motifs) are exposed (though Ser 152 and 154 are at the *C*-terminus) (Figure 7.1). The rest are either embedded inside the core (thus may not be solvent-exposed) or occur in the inter strand regions. This non-uniform distribution of these unique polar residues could be critical in designing an experimental strategy for mutagenesis.

Since mycoloylation of small proteins in *Corybacterium glutamicum* (e.g PorA) is catalysed by a mycoloylation transferase (Huc *et al*., 2010), it was envisaged that VapA could also be lipidated via a similar mechanism since lipoproteins and channel forming porins are non-covalently attached to the wall (Sutcliff *et al*., 2010). However, corynemycoloyl transferase did not catalyse the lipidation of VapA using extractable lipids from *Rhodococcus equi*, *Corynebacterium glutamicum* and *Mycobacterium bovis* BCG. In addition to catalysing the mycoloylation of small proteins (e.g. PorA), it has been discovered that Cmt also catalyses the lipidation of polypeptides in corynebacteria (Huc *et al*., 2013). Hence the failure of VapA to act as the final acyl acceptor in the *in vitro* model could be related to the fact that the protein was in a fully folded form and thus the mycoloylation site may not be available. This could be true if the protein is lipidated before export or translocation *via* the Sec machinery is complete.

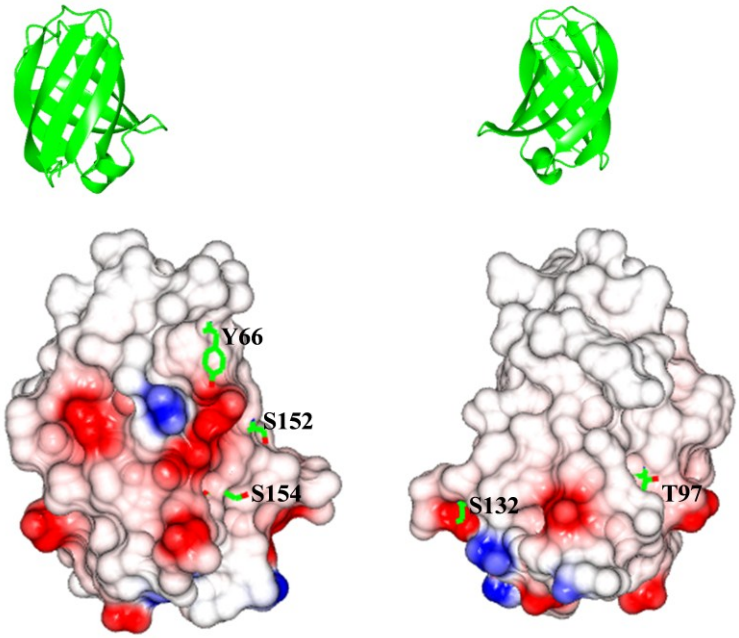


Figure 7.1 Electrostatic surface rendering of VapA model
Side chains of Ser, Thr, Tyr that are exposed are shown in fat bonds. The ribbon tracings (in green) show the orientation of the surfaces. The model was created using CCP4mg (McNicholas *et al*., 2011)

In summary, mycoloylation of proteins in *R. equi* is now established and the structural definition of the Vaps generated in this study and others (Whittingham *et al*., 2014, Geerds *et al*., 2014) provides a strong blue print for further study of Vap modification and function.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. (1990) Basic alignment search tool. **Journal of Molecular Biology** 215(3): 403-410.

Alvarez, H. M., Silva, R. A., Herrero, H., Hernandez, M. A. and Villalba, M. S. (2013) Metabolism of triacylglycerols in *Rhodococcus equi* species: insights from physiology and molecular genetics. **Journal of Molecular Biochemistry** 2(1): 69-78.

Baldwin, J. L., Bertone, J. J., Sommer, M. M., Bayha, R., Vaala, W., Cooper, W. L., Vanderwall, D. K. and Schalfer, D. H. (1992) *Rhodococcus equi* enteritis, colonic lymphadenitis, and peritonitis in three foals with nonresponsive *Rhodococcus equi* bronchopneumonia. **Equine Practice** 14(1): 15-18.

Barbey, C, Budin-Verneuil, A., Cauchard, S., Hartke, A., Laugier, C., Pichereau, V. and Petry, S. (2009) Proteomic analysis and immunogenicity of secreted proteins from *Rhodococcus equi* ATCC 33701. **Veterinary Microbiology** 135 (3-4): 334-345.

Barton, M. D. and Hughes, K. L. (1980) *Corynebacterium equi*: a review. **Veterinary Bulletin** 50: 65-80.

Barton, M. D. and Hughes, K. L. (1984) Ecology of *Rhodococcus equi*. **Veterinary Microbiology** 9(1): 65-76.

Belisle, J. T., Vissa, V. D., Sievert, T., Takayama, K., Brennan, P. J. and Besra, G. S. (1997) Role of major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. **Cell** 276(5317): 1420-1422.

Bendinger, B., Rijnarts, H. H. M., Altendorf, K. and Zehnder, A. J. B. (1993) Physicochemical cell surface and adhesive properties of corynebacteria related to the presence and chain length of mycolic acids. **Applied and Environmental Microbiology** 59 (11): 3973–3977.

Benoit, S., Benachour, A., Taouji, S., Auffray, Y. and Hartke, A. (2001) Induction of *vap* genes encoded by the virulence plasmid of *Rhodococcus equi* during acid tolerance response. **Research in Microbiology** 152(5): 439-449.

Benoit, S., Benachour, A., Taouji, S., Auffray, Y. and Hartke, A. (2002) H_2O_2, which causes macrophage- related stress, triggers induction of expression of virulence-associated plasmid determinants in *Rhodococcus equi*. **Infection and Immunity** 70 (7): 3768-3776.

Benoit, S., Taouji, S., Benachour, A. and Hartke, A. (2000) Resistance of *Rhodococcus equi* to acid pH. **International Journal of Food Microbiology** 55(1-3): 295-298.

Berleman, J. and Auer, M. (2013) The role of bacterial outer membrane vesicles for intra and interspecies delivery. **Environmental Microbiology** 15(2): 347-354.

Besenicar, M., Macek, P., Lakey, J. H. and Anderluh, G. (2006) Surface plasmon resonance in protein-membrane interactions. **Chemistry and Physics of Lipids** 141(1-2): 169-178.

Besra, G. S. and Brennan, P. J. (1997) The mycobacterial cell wall: biosynthesis of arabinogalactan and lipoarabinomannan. **Biochemical Society Transactions** 25(3): 845-850.

Bhatt, A., Molle, V., Besra, G. S., Jacobs, W. R. and Kremer, L. (2007) The *Mycobacterium tuberculosis* FAS-II condensing enzymes: their role in mycolic acid biosysthesis, acid-fastness, pathogenesis and in future drug development. **Molecular Microbiology** 64(6): 1442-1454.

Bidaud, P., Hebert, L., Barbey, C., Appourchaux, A-C., Torelli, R., Sanguinette, M., Laugier, C. and Petry, S. (2012) *Rhodococcus equi*'s extreme resistance to hydrogen peroxide is mainly conferred by one of its four catalase genes. **Plos ONE** 7(8): e42396.

Bordier, C. (1981) Phase separation of integral membrane proteins in Triton X-114 solution. **The Journal of Biological Chemistry** 256(4): 1604-1607.

Bordin, A. L., Suchodolski, J. S., Markel, M. E., Weaver, K. B., Steiner, J. M., Dowd, S. E., Pillai, S. and Cohen, N. D. (2013) Effects of administration of live or inactivated virulent *Rhodococcus equi* and age on the fecal microbiome of neonatal foals. **Plos One** 8(6): e66640.

Brand, S., Niehaus, K., Puhler, A. and Kalinowski, J. (2003) Identification and functional analysis of six mycolyltransferase genes of *Corynebacterium glutamicum* ATCC 13032: the genes *cop1*, *cmt1*, and *cmt2* can replace each other in the synthesis of trehalose dicorynomycolate, a component of the mycolic acid layer of the cell envelope. **Archives of Microbiology** 180(1): 33-44.

Brown, A. K., Sridharan, S., Dover, L. G., Sacchettini, J. C. and Besra, G. S. (2005) Probing the mechanism of the *Mycobacterium tuberculosis* *β*-Ketoacyl-Acyl Carrier Protein Synthase III *mt*FabH: Factors influencing catalysis and substrate specificity. **The Journal of Biological Chemistry** 280(37): 32539-32547.

Bryne, G. A., Boland, C. A., O′Connell, E. P. and Meijer, W. G. (2008) Differential mRNA stability of the *vapAICD* operon of the facultative intracellular pathogen *Rhodococcus equi*. **FEMS Microbiology Letters** 280(1): 89 – 94.

Bryne, G. A., Prescott, J. F., Palmer, G. H., Takai, S., Nicholson, V. M., Alperin, D. C. and Hines, S. A. (2001) Virulence plasmid of *Rhodococcus equi* contains inducible gene family encoding secreted proteins. **Infection and Immunity** 69(2): 650-656.

Bryne, G. A., Russel, D. A., Chen, X. and Meijer, W. G. (2007) Transcriptional regulation of the *virR* operon of the intracellular pathogen *Rhodococcus equi*. **Journal of Bacteriology** 189(14): 5082–5089

Buntain, S., Carter, C., Kuskie, K., Smith, J., Stepusin, R., Chaffin, M. K. and Takai, S. (2010) Frequency of *Rhodococcus equi* in feces of mares in Central Kentucky. **Journal of Equine Veterinary Science** 30(4): 191-195.

Burton, A. J., Giguere, S., Sturgill, T. L., Berghaus, L. J., Slovis, N. M., Whitman, J. L., Levering, C., Kuskie, K. R. and Cohen, N. D. (2013) Macrolide- and rifampin-resistant *Rhodococcus equi* on a horse breeding farm, Kentucky, USA. **Emerging Infectious Diseases** 19(2): 282-285.

Cardoso, S. A., Oliveira, A. F., Ruas, L. P., Trevisani, M. M., De Oliveira, L. L., Hanna, E. S., Roque-Barreira, M. C. and Soares, S. G. (2013) Nasal vaccination with attenuated *Salmonella* expression VapA: TLR2 activation is not essential for protection against *R. equi* infection. **Vaccine** 31(41): 4528-4535.

Chatterjee, D. and Khoo, K. H. (1998) Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. **Glycobiology** 8(2): 113-120.

Cohen, N. D., O'Conor, M. S., Chaffin, M. K. and Martens, R. J. (2005) Farm characteristics and management processes associated with development of *Rhodococcus equi* pneumonia in foals. **Journal of American Veterinary Medical Association** 226(3): 404-413.

Coulson, G. B., Agarwal, S. and Hondalus1, M. K. (2010) Characterization of the role of the pathogenicity island and *vapG* in the virulence of the intracellular actinomycete pathogen *Rhodococcus equi*. **Infection and Immunity** 78(8): 3323-3334.

Creuzenet, C., Durand, C. and Haertle, T. (1997) Interaction of α_s2- and β-casein signal peptides with DMPC and DMPG liposomes. **Peptides** 18(4): 463-472.

Darrah, P. A., Hondalus, M. K., Chen, Q., Ischiropoulos, H. and Mosser, D. M. (2000) Cooperation between reactive oxygen and nitrogen intermediates in killing of *Rhodococcus equi* by activated macrophages. **Infection and Immunity** 68(6): 3587-3593.

Dawson, T. R. M. Y., Horohov, D. W., Meijer, W. G. and Muscatello, G. (2010) Current understanding of the equine immune response to *Rhodococcus equi*. An immunological review of *R. equi* pneumonia. **Veterinary Immunology and Immunopathology** 135(1-2): 1-11.

Denham, E. L., Ward, P. N., and Leigh, J. A. (2009) In the absence of Lgt, lipoproteins are shed from *Streptococcus uberis* independently of Lsp. **Microbiology** 155(1): 134–141.

de Vagas, A. C., Monego, F., Gressler, L. T., Botton, S. A., Lazzari, A. M., da Costa, M. M., Ecco, R., Ribeiro, M. G., Lara, G. H. B. and Takai, S. (2013) Bronchopneumonia in wild boar (*Sus scrofa*) caused by *Rhodococcus equi* carrying the VapB type 8 plasmid. **BMC Research Notes** 6:111.

Desvaux, M., Hebraud, M., Talon, R. and Henderson, I. R. (2009) Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. **Trends in Microbiology** 17(4): 139-145.

Dias, M., Bhat, P., Chandrakar, S. and Pinto, H. (2013) *Rhodococcus equi*: A pathogen in immunocompetent patients. **Journal of Family Medicine and Primary Care** 2(3): 291-293.

Dong, X., Stothard, P., Forsythe, I. J. and Wishart D. S. (2004) PlasMapper: a web server for drawing and auto-annotating plasmid maps. **Nucleic Acids Research** 32: W660-W664.

Eikmanns, B. J., Kleinertz, E., Liebl, W. and Sahm, H. (1991) A family of Corynebacterium glutamicum/Escherichia coli shuttle vectors for gene cloning, controlled gene expression, and promoter probing. **Gene** 102(1), 93–98.

Elamin, A. A., Stehr, M., Oehlmann, W. and Singh, M. (2009) The mycolyltransferase 85A, a putative drug target of *Mycobacterium tuberculosis*: Development of a novel assay and quantification of glycolipid-status of the mycobacterial cell wall. **Journal of Microbacterial Methods** 79(3): 358–363.

Elissalde, G. S., Renshaw, H. W. and Walberg, J. A. (1980) *Corynebacterium equi*: an interhost review with emphasis on the foal. **Comparative Immunology, Microbiology and Infectious Diseases** 3(4): 433–445.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) Features and development of Coot. **Acta Crystallographica Section D Biological Crystallography** 66(4): 486-501.

Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K. and Welte, W. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. **Science** 282(5397): 2215-2220.

Fernandez-Mora, E., Polidori, M., Luhrmann, A., Schaible, U. E. and Haas, A. (2005) Maturation of *Rhodococcus equi*-containing vacuoles is arrested after completion of the early endosome stage. **Traffic** 6(8): 635-653.

Freire, E., Mayorga O. L. and Straume, M. (1990) Isothermal titration calorimetry. **Analytical Chemistry** 62(18): 950A-959A.

Fukuchi, S., Hosoda, K., Homma, K., Gojobori, T. and Nishikawa, K. (2011) Binary classification of protein molecules into intrinsically disordered and ordered segments. **BMC Structural Biology** 11: 29

Gande, R., Gibson, K. J. C., Brown, A. K., Krmbach, K., Dover, L. G., Sahm, H., Shioyama, S., Oikawa, T., Besra, G. S. and Eggeling, L. (2004) Acyl-CoA Carboxylases (*accD2* and *accD3*), together with a unique polyketide synthase (*Cg-pks*), are key to mycolic acid biosynthesis in *Corynebacterianeae* such as *Corynebacterium glutamicum* and *Mycobacterium tuberculosis*. **The Journal of Biological Chemistry** 279(43): 44847-44857.

Gao, L-Y., laval, F., Lawson, E. H., Groger, R. K., Woodruff, A., Morisaki, J. H., Cox, J. S., Daffe, M. and Brown, E. (2003) Requirement for *KasB* in *Mycobacterium* mycolic acid biosynthesis, cell wall impermeability and intracellular survival: implications for therapy. **Molecular Microbiology** 49(6): 1547-1563.

Geerds, C., Wohlmann, J., Haas, A., and Niemann, H. H. (2014) Structure of *Rhodococcus equi* virulence-associated protein B (VapB) reveals an eight-stranded antiparallel β-barrel consisting of two Greek-key motifs. **Acta Crystallographica Section F Structural Biology Communications** 70(7): 866-871.

Giguere, S. and Prescott, J. F. (1997) Clinical manifestations, diagnosis, treatment, and prevention of *Rhodococcus equi* infections in foals. **Veterinary Microbiology** 56(3-4): 313–334.

Giguere, S., Cohen, N. D., Keith Chaffin, M., Hines, S. A., Hondalus, M. K., Prescott, J. F. and Slovis, N. M. (2011) *Rhodococcus equi*: Clinical manifestations, virulence, and immunity. **Journal of Veterinary Internal Medicine** 25(6): 1221-1230.

Giguere, S., Hondalus, M. K., Yager, J. A., Darrah, P., Mosser, D. M. and Prescott, J. F. (1999) Role of the 85-kilobase plasmid and plasmid-encoded virulence-associated protein A in intracellular survival and virulence of Rhodococcus equi. **Infection and Immunity** 67(7): 3548-3557.

Giguère, S., Lee, E. A., Guldbech, K. M. and Berghaus, L. J. (2012) In vitro synergy, pharmacodynamics, and postantibiotic effect of 11 antimicrobial agents against *Rhodococcus equi*. **Veterinary Microbiology** 160(1-2): 207–213

González de la Vara, L. E. and Alfaro, B. L. (2009) Separation of membrane proteins according to their hydroparthy by serial phase partitioning with Triton X-114. **Analytical Biochemistry** 387(2): 280-286.

Goodfellow, M. (1987) The taxonomic status of *Rhodococcus equi*. **Veterinary Microbiology** 14(3): 205-209.

Goodfellow, M. and Alderson, G. (1977) The actinomycete-genus *Rhodococcus*: a home for the 'rhodochrous' complex. **Journal of General Microbiology** 100(1): 99–122.

Gouet, P., Robert, X and Courcelle, E. (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. **Nucleic Acids Research** 31(13): 3320–3323.

Grzegorzewicz, A. E., Pham, H., Gundi, V. A. K. B., Scherman, M. S., North, E. J., Hess, T., Jones, V., Gruppo, V., Born, S. E. M., Kordulakova, J., Chavadi, S. S., Morisseau, C., Lenaerts, A. J., Lee, R. E., McNeil, M. R. and Jackson, M. (2012) Inhibition of mycolic acid transport across the *Mycobacterium tuberculosis* plasma membrane. **Nature Chemical Biology** 8(4): 334-341.

Gurel, V., Lambert, K., Page, A. E., Loynachan, A. T., Huges, K., Timoney, J. F., Fettinger, M., Horohov, D. W. and McMichael, J. (2013) Streptolysin-O/antibiotics adjunct therapy modulates site-specific expression of extracellular matrix and inflammatory genes in lungs of *Rhodococcus equi* infected foals. **Veterinary Research Communications** 37(2): 145-154.

Harding, H. H. (2002) Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. **Acta Crystallographica Section D Biological Crystallography** 58(5): 872-874.

Hebert, L., Bidaud, P., Goux, D., Benachour, A., Laugier, C. and Petry, S. (2014) Study of lysozyme resistance in *Rhodococcus equi*. **Current Microbiology** 68(3): 352-357.

Hietala, S. K., Ardans, A. A. and Sansome, A. (1985) Detection of *Corynebacterium equi*-specific antibody in horses by enzyme-linked immunosorbent assay. **American Journal of Veterinary Research** 46(1): 13-15.

Hines, M. T., Paasch, K. M., Alperin, D. C., Palmer, G. H., Westhoff, N. C. and Hines, S. A. (2001) Immunity to *Rhodococcus equi*: antigen-specific recall responses in the lungs of adult horses. **Veterinary Immunology and Immunopathology** 79(1-2): 101–113.

Hines, S. A., Stone, D. M., Hines, M. T., Alperin, D. C., Knowles, D. P., Norton, L. K., Hamilton, M. J., Davis, W. C. and McGuire, T. C. (2003) Clearance of virulent but not avirulent *Rhodococcus equi* from the lungs of adult horses is associated with intracytoplasmic gamma interferon production by $CD4^+$ and

$CD8^+$ T Lymphocytes. **Clinical and Diagnostic Laboratory Immunology** 10(2): 208-215

Hondalus, M. K. (1997) Pathogenesis and virulence of *Rhodococcus equi*. **Veterinary Microbiology** 56(3-4): 257-268.

Hondalus, M. K. and Mosser, D. M. (1994) Survival and replication of *Rhodococcus equi* in Macrophages. **Infection and Immunity** 62(10): 4167-4175.

Hondalus, M. K., Diamond, M. S., Rosenthal, L. A., Springer, T. A. and Mosser, D. M. (1993) The intracellular bacterium *Rhodococcus equi* requires Mac-1 to bind to mammalian cells. **Infection and Immunity** 61(7): 2919-2929.

Horohov, D. W., Loynachan, A. T., Page, A. E., Hughes, K., Timoney, J. F., Fettinger, M., Hatch, T., Spaulding, J. G. and McMichael J. (2011) The use of streptolysin O (SLO) as an adjunct therapy for *Rhodococcus equi* pneumonia in foals. **Veterinary Microbiology** 154(1-2): 156-162.

Howlett, G. J., Minton, A. P. and Rivas, G. (2006) Analytical ultracentrifugation for the study of protein association and assembly. **Current Opinion in Chemical Biology** 10(5): 430-436.

Hsu, F-F, Soehl, K., Turk, J. and Haas, A. (2011) Characterization of mycolic acids from the pathogen *Rhodococcus equi* by tandem mass spectrometry with electrospray ionization. **Analytical Biochemistry** 409(1): 112-122.

Huang, H. and Sarai, A. (2012) Analysis of the relationships between evolutionary, thermodynamics, and the functions of intrinsically disordered proteins/regions. **Computational Biology and Chemistry** 41: 51-57.

Huc, E., de Sousa-D'Auria, C., de la Sierra-Gallay, I. L., Salmeron, C., van Tilbeurgh, H., Bayan, N., Houssin, C., Daffe, M. and Tropis, M. (2013) Identification of a mycoloyl transferase selectively involved in O-acylation of polypeptides in Corynebacteriales. **Journal of Bacteriology** 195(18): 4121-4128.

Huc, E., Meniche, X., Benz, R. Bayan, N., Ghazi, A., Tropis, M. and Daffe, M. (2010) *O*-Mycoloylated proteins from *Corynebacterium* - An unprecedented post-translational modification in bacteria. **The Journal of Biological Chemistry** 285(29): 21908-21912

Hughes, K. L. and Sulaiman, I. (1987) The ecology of *Rhodococcus equi* and physicochemical influences on growth. **Veterinary Microbiology** 14(3): 241-250.

Hutchings, M. H., Palmer, T., Harrington, D. J. and Sutcliffe, I. C. (2009) Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. **Trends in Microbiology** 17(1): 13-21.

Jacks, S., Giguere, S. and Prescott, J. F. (2007) In vivo expression of and cell-mediated immune responses to the plasmid-encoded virulence-associated proteins of *Rhodococcus equi* in foals. **Clinical and Vaccine Immunology** 14(4): 369-374.

Jain, S., Bloom, B. R. and Hondalus, M. K. (2003) Deletion of *vapA* encoding virulence associated protein-A attenuates the intracellular actinomycete *Rhodococcus equi*. **Molecular Microbiology** 50(1): 115-128.

Jarlier, V. and Nikaido, H. (1994) Mycobacterial cell wall: Structure and role in natural resistance to antibiotics. **FEMS Microbiology Letters** 123(1-2): 11-18.

Jones, A. L., Sutcliffe, I. C. and Goodfellow, M. (2013a) *Prescottia equi* gen. nov., comb. Nov.: a new home for an old pathogen. **Antonie van Leeuwenhoek** 103(3): 655-671.

Jones, A. L., Sutcliffe, I. C. and Goodfellow, M. (2013b) Proposal to replace the illegitimate genus name *Prescottia* Jones et al. 2013 with the genus name Prescottella gen. nov. and to replace the illegitimate combination Prescottia equi Jones et al. 2013 with *Prescottella equi* comb. nov. **Antonie van Leeuwenhoek** 103(6): 1405-1407.

Joshi, A. B., Sawai, M., Kearney, W. R. and Kirsch, L. E. (2005) Studies on the mechanism of aspartic acid cleavage and glutamine deamidation in the acidic degradation of glucagon. **Journal of Pharmaceutical Sciences** 94(9): 1912–1927.

Kabsch, W. (2010). XDS. **Acta Crystallographica Section D Biological Crystallography** 66(2) 125-132.

Kacem, R., Sousa-D'Auria, C. D., Tropis, M., Chami, M., Gounon, P., Leblon, G., Houssin, C. and Daffe, M. (2004). Importance of mycoloyltransferases on the physiology of *Corynebacterium glutamicum*. **Microbiology** 150(1): 73 – 84.

Kai, M., Fujita, Y., Maeda, Y., Nakataa, N., Izumi, S., Yano, I. and Makino, M. (2007) Identification of trehalose dimycolate (cord factor) in *Mycobacterium leprae*. **FEBS Letters** 581(18): 3345-3350.

Kanaly, S. T., Hines, S. A. and Palmer, G. H. (1995) Cytokine modulation alters pulmonary clearance of *Rhodococcus equi* and development of granulomatous pneumonia. **Infection and Immunity** 63(8): 3037-3041.

Kaufmann, S. H. and Hess, J. (1999) Impact of intracellular location of and antigen display by intracellular bacteria: implications for vaccine development. **Immunology Letters** 65(1-2): 81-84.

Kaura, Y. K. and Mutimer, M. D. (1987) Biochemical and serological characteristics of Indian strains of *Rhodococcus equi* (*Corynebacterium equi*). **Indian Journal of Microbiology** 27: 39-42.

Kittelberger, R., Hansen, M. F., Hilbink, F., de Lisle, G. W. and Cloeckaert, A. (1995) Selective extraction of bacterial macromolecules by temperature-induced phase separation in Triton X-114. **Journal of Microbiology Methods** 24(1): 81-92.

Kovacs-Simon, A., Titball, R. W. and Michell, S. L. (2011) Lipoproteins of Bacterial Pathogens. **Infection and Immunity** 79(2): 548-561.

Kozłowska, M., Tarczewska, A., Jakob, M., Szpotkowki, K., Wojtas, M., Rymarczyk, G. and Ozyhar, A. (2014) Calponin-Like Chd64 is partly disordered. **Plos One** 9(5): e96809.

Kremer, L., Maughan, W. N., Wilson, R. A., Dover, L. G. and Besra, G. S. (2002) The *M. tuberculosis* antigen 85 complex and mycolyltransferase activity. **Letters in Microbiology** 34(4): 233-237.

Kuehn, M. J. and Kesty, N. C. (2005) Bacterial outer membrane vesicles and the host- pathogen interaction. **Genes and Development** 19: 2645-2655.

Ladbury, J. E. and Chowdhry, B. Z. (1996) Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. **Chemistry and Biology** 3(10): 391-801.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature** 227: 680-685.

Lavental, I., Grzybek, M. and Simons, K. (2010) Greasing their way: Lipid modification determine protein association with membrane rafts. **Biochemistry** 49: 6305-6316.

Lee, E-Y., Choi, D-Y., Kim D-K., Kim, J-W., Park J. O., Kim, S., Kim., S-H., Desiderio, D. M., Kim, Y-K., Kim, K-P. and Gho, Y. S. (2009) Gram-positive bacteria produce membrane vesicles: Proteomics-based characterisation of *Staphylococcus aureus*-derived membrane vesicles. **Proteomics** 9(24): 5425-5436.

Lee, R. E., Armour J. W., Takayama, K., Brennan P. J. and Besra, G. S. (1997) Mycolic acid biosynthesis: definition and targeting of the Claisen condensation step. **Biochimica et Biophysica Acta** 1346(3): 275-284.

Letek, M., Gonza´ lez, P., MacArthur, I., Rodrı´guez, H., Freeman, T. C., Valero-Rello, A., Blanco, M., Buckley, T., Cherevach, I., Fahey, R., Hapeshi, A., Holdstock, J., Leadon, D., Navas, J., Ocampo, A., Quail, M. A., Sanders, M., Scortti, M. M., Prescott, J. F., Fogarty, U., Meijer, W. G., Parkhill, J., Bentley,

S. D. and Va´zquez-Boland, J. A. (2010) The genome of a Pathogenic *Rhodococcus*: Cooptive virulence underpinned by key gene acquisitions. **Plos Genetics** 6(9): e1001145.

Letek, M., Ocampo-Sosa, A. A., Sanders, M., Fogarty, U., Buckley, T., Leadon, D. P., Gonza´lez, P., Scortti, M., Meijer, W. G., Parkhill, J., Bentley, S. and Va´zquez-Boland, J. A. (2008) Evolution of the *Rhodococcus equi vap* pathogenicity island seen through comparison of host-associated *vapA* and *vapB* virulence plasmids. **Journal of Bacteriology** 190(17): 5797-5805.

Lichtinger, T., Burkovski, A., Niederweis, M., Kramer, R. and Benz, R. (1998) Biochemical and biophysical characterization of the cell wall channel of *Corynebacterium glutamicum*: The channel is formed by a low molecular mass subunit. **Biochemistry** 37: 15024-15032.

Lichtinger, T., Heym, B., Maier, E., Eichner, H., Cole, S. T. and Benz, R. (1999) Evidence for a small anion-selective channel in the cell wall of *Mycobacterium bovis* BCG besides a wide cation-selective pore. **FEBS Letters** 454(3): 349-355.

Lichtinger, T., Reiss, G. and Benz, R. (2000) Biochemical identification and biophysical characterization of a channel-forming protein from *Rhodococcus erythropolis*. **Journal of Bacteriology** 182(3): 764-770.

Linder, R. and Bernheimer, A. W. (1997) Oxidation of macrophage membrane cholesterol by intracellular *Rhodococcus equi*. **Veterinary Microbiology** 56(3-4): 269-276.

Lohmann, K. L., Lopez, A. M., Manning, S. T., Marques, F. J., Brownlie, R. B., Allen, A. L., Sangster, A. E., Mutwiri, G., Gerdts, V., Potter, A. and Townsend, H. G. G. (2013) Failure of a VapA/CpG oligodeoxynucleotide vaccine to protect foals against experimental *Rhocococcus equi* pneumonia despite induction of VapA-specific antibody and interferon-γ response. **Canadian Journal of Veterinary Research** 77(3): 161-169.

Lopez, A. M., Hines, M. T., Palmer, G. H., Alperin, D. C. and Hines, S. A. (2002) Identification of pulmonary T-lymphocytes and serum antibody isotype responses associated with protection against *Rhodococcus equi*. **Clinical and Diagnostic Laboratory Immunology** 9(6): 1270-1276.

MacArthur, I., Parreira, V. R., Lepp, D., Mutharia, L. M., Vazquez-Boland, J. A. and Prescott, J. F. (2011) The sensor kinase MprB is required for *Rhodococcus equi* virulence. **Veterinary Microbiology** 147(1-2): 133-141.

Mafakhari, S., Barcena-Uribarri, A., Abdali, N., Jones, A. L., Sutcliffe, I. C. and Benz, R. (2014) Discovery of a cell wall porin in the mycolic-acid-containing actinomyces *Dietzia maris* DSM 43672. **FEBS Journal** 281(8): 2030-2041.

Magnusson, H. (1923) Spezifische infektiose Pneumonie beim Fohlen. Ein neuer Eiterreger beim Pferd Archiv Wissenschaft Praktit-Tierheilkunde 50: 22-38.

Makrai, L., Takai, S., Tamura, M., Tsukamoto, A., Sekimoto, R., Sasaki, Y., Kakuda, T., Varga, J., Fodor, L., Solymosi, N. and Major, A. (2002) Characterization of virulence plasmid types in *Rhodococcus equi* isolates from foals, pigs, humans and soil in Hungary. **Veterinary Microbiology** 88(4): 377-384.

Marraffini, L. A., DeDent, A. C. and Olaf, S. (2006) Sortases ad the art of anchoring proteins to the envelopes of Gram-positive bacteria. **Microbiology and Molecular Biology Reviews** 70(1): 192-221.

Marrakchi, H., Laneelle, M-A and Daffe, M. (2014) Mycolic acids: Structure, biosynthesis, and beyond. **Chemistry and Biology** 21(1): 67-85.

McNeil, M. M. and Brown, J. M. (1994) The medically important aerobic actinomycetes: epidemiology and microbiology. **Clinical Microbiology Reviews** 7(3): 357-417.

McNicholas, S., Potterton, E., Wilson, K. S. and Noble, M. E. (2011) Presenting your structures: the CCP4mg molecular-graphics software. **Acta Crystallographica Section D Biological Crystallography** 67(4): 386-394.

Meeuse, J. J., Sprenger, H. G., van Assen, S., Leduc, D., Daenen, S. M. G. J., Arends, J. P. and van der Werf, T. S. (2007) *Rhodococcus equi* infection after Alemtuzumab therapy for T-cell prolymphocytic leukemia. **Emerging Infectious Diseases** 13(12): 1942-1943.

Meijer, W. G. and Prescott, J. F. (2004) *Rhodococcus equi*. **Veterinary Research** 35(4): 383-396.

Minnikin, D. E. (1982) Lipids: complex lipids, their chemistry, biosynthesis and roles: In: Ratledge, C. and Stanford, J. L. Eds. **The Biology of the Mycobacteria; Physiology, identification and classification.** Academic Press, London. Vol 1, pp 95-184

Miranda-CasoLuengo, R., Coulson, G. B., Miranda-CasoLuengo, A., Vazquez-Boland, J. A., Hondalus, M. K. and Meijer, W. G. (2012) The hydroxamate siderophore rhequichelin is required for virulence of the pathogenic actinomycete *Rhodococcus equi*. **Infection and Immunity** 80(12): 4106-4114.

Miranda-CasoLuengo, R., Miranda-CasoLuengo, A. A., O'Connell, E. P., Fahey, R. J., Boland, C. A., Vazquez-Boland, J. A. and Meijer, W. G. (2011) The *vapA* co-expressed virulence plasmid gene *vcgB* (*orf10*) of the intracellular actinomycete *Rhodococcus equi*. **Microbiology** 157(8): 2357-2368.

Mizianty, M. and Kurgan, L. (2011) Improved identification of outer membrane beta barrel proteins using primary sequence, predicted secondary structure, and evolutionary information. **Proteins** 79(1): 294-303.

Moreira, L. G. A., Pereira, L. C., Drummond, P. R. and De Mesquita, J. F. (2013) Structural and functional analysis of human SOD1 in amyotrophic lateral sclerosis**. Plos One** 8(12): e81979.

Mosser, D. M. and Hondalus, M. K. (1996) *Rhodococcus equi*: an emerging opportunistic pathogen. **Trends in Microbiology** 4(1): 29-33.

Mukherjee, S., Hao Y-H. and Orth, K. (2007) A newly discovered post-translational modification-the acetylation of serine and threonine residues. **Trends in Biochemical Sciences** 32(5): 210-216.

Munoz, M., Luquin, M., Garcia-Barcelo, M., Julian, E., Ausina, V. and Laneelle, M. A. (1997) Distribution of surface-exposed antigenic glycolipids in recent clinical isolates of *Mycobacterium tuberculosis*. **Research in Microbiology** 148(5): 405-412.

Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. **Acta Crystallographica Section D Biological Crystallography** 67(4): 355-367.

Muscatello, G. (2012) *Rhodococcus equi* pneumonia in the foal-Part 1: Pathogenesis and epidemiology. **The Veterinary Journal** 192(1): 20-26.

Nath, S. R., Mathew, A. P., Mohan, A. and Anila, K. R. (2013) *Rhodococcus equi* granulomatous mastitis in an immunocompetent patient**. Journal of Medical Microbiology** 62(8): 1253-1255.

Navarre, W. W. and Schneewind, O. (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. **Microbiology and Molecular Biology Reviews** 63(1): 174-229.

Nelson, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. **Protein Engineering** 10(1): 1–6.

Nigou, J., Gilleron, M. and Puzo, G. (2003) Lipoarabinomannans: from structure to biosynthesis. **Biochimie** 85(1-2): 153-166.

Nikaido, H. (1994) Porins and specific diffusion channels in bacterial outer membranes. **The Journal of Biological Chemistry** 269(6): 3905-3908

Nordmann, P., Kerestedjian, J. J. and Ronco, E. (1992) Role of T-Lymphocyte subsets in *Rhodococcus equi* Infection. **Infection and Immunity** 60(7): 2748-2752

Oldfield, C., Bonella, H., Renwick, L., Dodson, H. I., Alderson, G. and Goodfellow, M. (2004) Rapid determination of *vapA/vapB* genotype in *Rhodococcus equi* using a differential polymerase chain reaction method. **Antonie van Leeuwenhoek** 85(4): 317-326.

Ota, M., Koike, R., Amemiya, T., Tenno, T., Romero, P. R., Hiroaki, H., Dunker, A. K. and Fukuchi, S. (2013) An assignment of intrinsically disordered regions of proteins based on NMR structures. **Journal of Structural Biology** 181(1): 29-36.

Patching, S. G. (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. **Biochimica et Biophysica Acta** 1838(1): 43-55.

Patton, K. M., McGuire, T. C., Hines, M. T., Mealey, R. H. and Hines, S. A. (2005) *Rhodococcus equi*-specific cytotoxic T-lymphocytes in immune horses and development in asymptomatic foals. **Infection and Immunity** 73(4): 2083–2093.

Pauptit, R. A., Schirmer, T., Jansonius, J. N., Rosenbusch, J. P., Parker, M. W., Tucker, A. D., Tsernoglou, D., Weiss, M. S. and Schultz, G. E. (1991) A common channel-forming motif in evolutionarily distant porins. **Journal of Structural Biology** 107(2):136-145.

Perez, M. G. V., Vassilev, T. and Kemmerly, S. A. (2002) *Rhodococcus equi* infection in transplant recipients: a case of mistaken identity and review of the literature. **Transplant Infectious Disease** 4(1): 52-56.

Petersen, T. N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods** 8(10): 785-786.

Phumoonna, T., Muscatello, G., Chicken, C., Gilkerson, J. R., Browning, G. F., Barton, M. D. and Heuzenroeder, M. W. (2006) Clinical evaluation of a Peptide-ELISA based upon N-terminal B-cell epitope of the VapA protein for diagnosis of *Rhodococcus equi* pneumonia in foals. **Journal of Veterinary Medicine** 53(3): 126–132.

Polidori, M. and Haas, A. (2006) VapI, a new member of the *Rhodococcus equi* Vap family. **Antonie van Leeuwenhoek** 90(3): 299-304.

Poolkhet, C., Chumsing, S., Wajjwalku, W., Minato, C., Otsu, Y. and Takai, S. (2010) Plasmid profiles and prevalence of intermediately virulent *Rhodococcus equi* from pigs in Nakhonpathom province, Thailand: Identification of a new

variant of the 70-kb virulent plasmid, type 18. **Veterinary Medicine International.** DOI 4061/2010/491624.

Portevin, D., de Sousa-D′Auria, C., Housin, C., Grimaldi, C., Chami, M., Daffe, M. and Guilhot, C. (2004) A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. **Proceedings of the National Academy of Sciences** 101(1): 314-319.

Porto, A. C. R. C., Petrello, M. S., Hoge, A. Y. A., Ambrozio, G. R.,Massoco, C. O. and Oliveira, C. A. (2014) Equine colostrum: more than IgG. **Journal of Equine Veterinary Science** 34(1): 239.

Prescott, J. F. (1987) Epidemiology of *Rhodococcus equi* infection in horses. **Veterinary Microbiology** 14(3): 211-214.

Prescott, J. F. (1991) *Rhodococcus equi*: An Animal and Human pathogen. **Clinical Microbiology Reviews** 4(1): 20-34.

Puech, V., Chami, M., Lemassu, A., Laneelle, M-A., Schiffler, B., Gounon, P., Bayan, N., Benz, R. and Daffe, M. (2001) Structure of the cell envelope of corynebacteria: importance of the non-covalently bound lipids in the formation of the cell wall permeability barrier and fracture plane. **Microbiology** 147(5): 1365-1382.

Qin, J. and Gronenborn, A. M. (2014) Weak protein complexes: challenging to study but essential for life. **FEBS Journal** 281(8): 1948-1949.

Raad R. B., Meniche, X., de Sousa-d'Auria, C., Chami, M., Salmeron, C., Tropis, M., Labarre, C., Daffe, M., Houssin, C. and Bayan, N. (2010) A deficiency in arabinogalactan biosynthesis affects *Corynebacterium glutamicum* mycolate outer membrane stability. **Journal of Bacteriology** 192 (11): 2691-2700.

Raman, K., Rajagopalan, P. and Chandra, N. (2005) Flux balance analysis of mycolic acid pathway: Target for anti-tubercular drugs. **Plos Computational Biology** 1(5): 0349-0358.

Rath, P., Demange, P., Saurel, O., Tropis, M., Daffe, M., Dotsch, V., Ghazi, A., Bernhard, F. and Milon, A. (2011) Functional expression of the PorAH channel from *Corynebacterium glutamicum* in cell-free expression Systems; implications for the role of the naturally occurring mycolic acid modification. **The Journal of Biological Chemistry** 286(37): 32525-32532.

Rath, P., Saurel, O., Tropis, M., Daffe, M., Demange, P. and Milon, A. (2013) NMR localization of the O-mycoloylation on PorH, a channel forming peptide from *Corynebacterium glutamicum*. **FEBS Letters** 587(22): 3687-3691.

Ratledge, C. and Dover, L. G. (2000) Iron metabolism in pathogenic bacteria. **Annual Reviews of Microbiology** 54: 881-941.

Ren, J. and Prescott, J. F. (2003) Analysis of virulence plasmid gene expression of intra-macrophage and *in vitro* grown *Rhodococcus equi* ATCC 33701. **Veterinary Microbiology** 94(2): 167-182.

Reuss, S. M., Chaffin, M. K. and Cohen, N. D. (2009) Extrapulmonary disorders associated with *Rhodococcus equi* infections in foals: 150 cases (1987-2007). **Journal of the American Veterinary Medical Association** 235(7): 855-863.

Ribero, M. G., Takai, S., Guazzelli, A., Lara, G. H. B., da Silva, A. V., Fernandes, M. C., Condas, L. A. Z., Siqueira, A. K. and Salerno, T. (2011) Virulence genes and plasmid profiles in *Rhodococcus equi* isolates from pigs and wild boars (*Sus scrofa*) in Brazil. **Research in Veterinary Science** 91(3): 478-481.

Rich, R. L. and Myszka, D. G. (2000) Advances in surface plasmon resonance biosensor analysis. **Current Opinion in Biotechnology** 11(1): 54-61.

Riess, F. G. and Benz, R. (2000) Discovery of a novel channel-forming protein in the cell wall of the non-pathogenic *Nocardia corynebacteroides*. **Biochimica et Biophysica Acta** 1509(1-2): 485-495.

Riess, F. G., Elflein, M., Benk, M., Schiffler, B., Benz, R., Garton, N. and Sutcliffe, I. (2003) The cell wall of the pathogenic bacterium *Rhodococcus equi* contains two channel-forming proteins with different properties. **Journal of Bacteriology** 185(9): 2952-2960.

Rigel, N. W. and Silhavy, T. J. (2012) Making a beta-barrel: assembly of outer membrane proteins in Gram-negative bacteria. **Current Opinion in Microbiology** 15(2): 189-193.

Rzewuska, M., Wilkowski, L., Cisek, A. A., Stefanska, I., Chrobak, D., Stefaniuk, E., Kizerwetter-Swida, M. and Takai, S. (2014) Characterization of *Rhodococcus equi* isolates from submaxillary lymgh nodes of wild boars (*Sus scrofa*), red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*). **Veterinary Microbiology** 172(1-2): 272-278.

Sali, A. (1995) Comparative protein structure modelling by satisfaction of spatial restraints. **Molecular Medicine Today** 1(6): 270-277.

Salifu, S. P., Campbell Casy, S. A. and Foley, S. (2013) Isolation and characterization of soilborne virulent bacteriophages infecting the pathogen *Rhodococcus equi*. **Journal of Aplied Microbiology** 114(6): 1625-1633.

Sambrook, J. and Russell, D. W. (2001) Preparation and transformation of competent *E. coli* using calcium chloride. In; **Molecular Cloning: A Laboratory Manual**. Third Edition pp 1.116- 1.118.

Sander, P., Rezwan, M., Walker, B., Rampini, S. K., Kroppenstedt, R. M., Ehlers, S., Keller, C., Keeble, J. R., Hagemeier, M., Colston, M. J., Springer, B. and Bottger, E. C. (2004) Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*. **Molecular Microbiology** 52(6): 1543-1552.

Sangal, V., Jones, A. L., Goodfellow, M., Sutcliffe, I. C. and Hoskisson, P. A. (2014) Comparative genomic analyses reveal a lack of a substantial signature of host adaptation in *Rhodococcus equi* (‘*Prescottella equi*’). **Pathogens and Diseases** 71(3): 352-356.

Siroy, A., Mailaender, C., Harder, D., Koerber, S., Wolschendorf, F., Danilchanka, O., Wang, Y., Heinz, C. and Niederweis, M. (2008) Rv1698 of *Mycobacterium tuberculosis* represents a new class of channel-forming outer membrane proteins. **The Journal of Biological Chemistry** 283(26): 17827-17837.

Smither, S. J., Hill, J., van Baar, B. L. M., Hulst, A. G., de Jong, A. L. and Titball, R. W. (2007) Identification of outer membrane proteins of *Yersinia pestis* through biotinylation. **Journal of Microbiology Methods** 68(1): 26-31.

Stahelin, R. V. (2013) Surface plasmon resonance: a useful technique for cell biologists to characterize biomolecular interactions. **Molecular Biology of the Cell** 24(7): 883-886.

Sutcliffe, I. C. (1998) Cell envelope composition and organisation in the genus *Rhodococcus*. **Antonie van Leeuwenhoek** 74(1-3): 49-58.

Sutcliffe, I. C. and Harrington, D. J. (2002) Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. **Microbiology** 148(7): 2065-2077.

Sutcliffe, I. C. and Russell, R. R. B. (1995) Lipoproteins of Gram-positive bacteria. **Journal of Bacteriology** 177(5): 1123-1128.

Sutcliffe, I., Brown A. K. and Dover L.G. (2010) The Rhodococcal Cell Envelope: Composition, Organisation and Biosynthesis. In: Alvarez HM, ed. Biology of *Rhodococcus*, **Microbiology Monographs** 16. Berlin Heidelberg: Spinger-Verlag, pp 29 – 71.

Sydor, T., von Bergen, K., Hsu, F-F, Huth, G., Holst, O., Wohlmann, J., Becken, U., Dykstr, T., Sohl, K., Lindner, B., Prescott, J. F., Schalble, U. E., Utermohlen, O. and Hass, A. (2013) Diversion of phagosome trafficking by pathogenic *Rhodococcus equi* depends in mycolic acid chain length. **Cellular Microbiology** 15(3): 458-473.

Tahlan, K., Wilson, R., Kastrinsky, D. B., Arora, K., Fischer, E., Barnes, S. W., Walker, J. R., Alland, D., Barry, C. E. and Boshoff, H. (2012) SQ109 targets MmpL3, a membrane transporter of trehalose monomycolate involved in

mycolic acid donation to the cell wall core of *Mycobacterium tuberculosis*. **Antimicrobial Agents and Chemotherapy** 56(4): 1797-1809.

Takai, S. (1997) Epidemiology of *Rhodococcus equi* infections: A review. **Veterinary Microbiology** 56(3-4): 167-176.

Takai, S., Fukunaga, N., Ochiai, S., and Imai, Y., Sasaki, Y., Tsubaki, S. and Sekizaki, T. (1996) Identification of intermediately virulent *Rhodococcus equi* isolates from pigs**. Journal of Clinical Microbiology** 34(4): 1034-1037.

Takai, S., Hines, S. A., Sekizaki, T., Nicholson, V. M., Alperin, D. A., Osaki, M., Takamatsu, D., Nakamura, M., Suzuki, K., Ogino, N., Kakuda, T., Dan, H. and Prescott, J. F. (2000) DNA Sequence and comparison of virulence plasmids from *Rhodococcus equi* ATCC 33701 and 103. **Infection and Immunity** 68(12): 6840-6847.

Takai, S., Martens, R. J., Julian, A., Ribeiro, M. G., de Farias, M. R., Sasaki, Y., Inuzuka, K., Kakuda, T., Tsubaki, S. and Prescott, J. F. (2003) Virulence of *Rhodococcus equi* isolated from cats and dogs. **Journal of Clinical Microbiology** 41(9): 4468-4470.

Tan, C., Prescott, J. F., Patterson, M. C. and Nicholson, V. M. (1995) Molecular characterization of a lipid-modified virulence-associated protein of *Rhodococcus equi* and its potential in protective immunity. **Canadian Journal of Veterinary Research** 59(1): 51-59.

Toyooka, K., Takai, S. and Kirikae, T. (2005) *Rhodococcus equi* can survive a phagolysosomal environment in macrophages by suppressing acidification of the phagolysosome. **Journal of Medical Microbiology** 54(11): 1007-1015.

Tudors, A. J. and Schasfoort, R. B. M. (2008) Introduction to surface plasmon resonance. In: Schasfoort, R. B. M. and Tudor, A. J. (Eds); **Handbook of Surface Plasmon Resonance**. The Royal Society of Chemistry, pp 1-14.

Ueda, S., Fujiwara, N., Naka, T., Sakaguchic, I., Ozekia, Y., Yanod, I., Kasamae, T. and Kobayashia, K. (2001) Structure–activity relationship of mycoloyl glycolipids derived from *Rhodococcus* sp. 4306. **Microbial Pathogenesis** 30(2): 91-99.

Usha, V., Dover, L. G., Roper, D. L., Lloyd, A.J. and Besra, G. S. (2006) Use of a codon alteration strategy in a novel approach to cloning the *Mycobacterium tuberculosis* diaminopimelic acid epimerase. **FEMS Microbiology Letters** 262(1): 39-47.

Uversky, V. N. (2011) Intrinsically disordered proteins from A to Z. **The International Journal of Biochemistry and Cell Biology** 43(8): 1090-1103.

Uversky, V. N. (2013) Unusual biophysics of intrinsically disordered proteins. **Biochimica et Biophysica Acta** 1834(5): 932-951.

Vagin, A. and Teplykov, A. (2010). Molecular replacement with MOLREP. **Acta Crystallographica Section D Biological Crystallography** 66 (1): 22-25.

Van der Geize, R., Grommen, A. F. W., Hessels, G. I., Jacobs, A. A. C. and Dijkhuizen, L. (2011) The steroid catabolic pathway of the intracellular pathogen *Rhodococcus equi* is important for pathogenesis and a target for vaccine development. **Plos Pathogens** 7(8): e1002181.

Van der Rest, M. E., Lange, C., and Molenaar, D. (1999) A heat shock following electroporation induces highly efficient transformation of *Corynebacterium glutamicum* with xenogeneic plasmid DNA. **Applied Microbiology and Biotechnology** 52(4): 541-545.

Van Roosmalen, M. L., Geukens, N., Jongbloed, J. D. H., Tjalsma, H., Dubois, J-Y. F., Bron, S, van Dijl, J. M. and Anne, J. (2004) Type I signal peptidases of gram-positive bacteria. **Biochimie et Biophysica Acta** 1694(1-3): 279-297.

Van Wely, K. H. M., Swaving, J., Freudl, R. and Driessen, A. J. M. (2001) Translocation of proteins across the cell envelope of gram-positive bacteria. **FEMS Microbiology Reviews** 25(4): 437-454.

Vazquez-Boland, J. A., Giguere, S., Hapeshi, A., MacArthur, I., Anastasi, E. and Valero-Rello, A. (2013) *Rhodococcus equi*: The many facets of a pathogenic actinomycete. Veterinary Microbiology. **Veterinary Microbiology** 167(1-2): 9-33.

Vazquez-Boland, J. A., Letek, M., Valero, A., Gonzalez, P., Scortti, M. and Fogarty, U. (2010) *Rhodococcus equi* and its pathogenic mechanisms. In: Alvarez HM, ed. **Biology of Rhodococcus**, Microbiology Mongraphs 16. Berlin Heidelberg: Spinger-Verlag. pp 331-359.

Velazquez Benito, A., Juste Tejero, C., Perez Lazaro, C. and Santos Lasaosa, S. (2013) Cerebral abscess due to *Rhodococcus equi* with pseodotumour immunocompetent patient. **Neurologia** 28(8): 522-524.

Venner, M., Astheimer, K., Lammer, M. and Giguere, S. (2013) Efficacy of mass antimicrobial treatment of foals with subclinical pulmonary abscesses associated with *Rhodococcus equi*. **Journal of Veterinary Internal Medicine** 27(1):171–176.

Vera-Cabrera, L., Ortiz-Lopez, R., Elizondo-Gonzalez, R. and Ocampo-Candiani, J. (2013) Complete genome sequence analysis of *Nocardia brasiliensis* HUJEG-1 reveals a saprobic lifestyle and the genes needed for human pathogenesis. **Plos One** 8(6): e65425.

Vercellone, A., Nigou, J. and Puzo, G. (1998) Relationships between the structure and the roles of lipoarabinomannans and related glycoconjugates in tuberculosis pathogenesis. **Frontiers in Bioscience** 3: e149-e163.

Verschoor, J. A., Baird, M. S. and Grooten, J. (2012) Towards understanding the functional diversity of the cell wall mycolic acids of *Mycobacterium tuberculosis***. Progress in Lipid Research** 51(4): 325-339.

Villalba, M. S., Hernandez, M. A., Silva, R. A. and Alvarez, H. M. (2013) Genome sequences of triacylglycerol metabolism in *Rhodococcus* as a platform for comparative genomics. **Journal of Molecular Biochemistry** 2: 94-105.

Volkin, D. B., Mach, H. and Middaugh, C. R. (1997) Degradative covalent reactions important to protein stability. **Molecular Biotechnology** 8(2): 105–122.

Von Bargen, K. and Haas, A. (2009) Molecular and infection biology of the horse pathogen *Rhodococcus equi*. **FEMS Microbiology Reviews** 33(5): 870-891.

Wang, X., Coulson, G. B., Miranda-CasoLuengo, A. A., Miranda-CasoLuengo, R., Hondalus, M. K. and Meijer, W. G. (2014) IcgA is a virulence factor of *Rhodococcus equi* that modulates intracellular growth. **Infection and Immunity** 82(5): 1793-1800.

Webb, B. and Sali, A. (2014) Protein structure modelling with Modeller. **Methods in Molecular Biology** 1137:1-15.

Weickert, M. J., Doherty, D. H., Best, E. A. and Olins, P. O. (1996) Optimisation of heterologous protein production in *Escherichia coli*. **Current Opinion in Biotechnology** 7 (5): 494-499.

Whittingham, J. L., Blagova, E. V., Finn, C. E., Luo, H., Miranda-CasoLuengo, R., Turkenburg, J. P., Leech, A. P., Walton, P. H., Murzin, A. G., Meijer, W. G. and Wilkinson, A. T. (2014) Structure of the virulence-associated protein VapD from the intracellular pathogen *Rhodococcus equi*. **Acta Crystallographica Section D Biological Crystallography** 70(8): 2139-2151.

Wiederstein, M. and Sippl M. J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acid Research** 35(suppl 2): W407–W410.

Wimley, W. C. (2003) The versatile beta-barrel membrane protein. **Current Opinion in Strcutural Biology** 13 (4): 404-411.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J.,

Vagin, A. A. and Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. **Acta Crystallographica Section D Biological Crystallography** 67(4): 235-242.

Yager, J. A. (1987) The pathogenesis of *Rhodococcus equi* pneumonia in foals. **Veterinary Microbiology** 14(3): 225-232.

Yamshchikov, A. V., Schuetz, A. and Marshall Lyon, G. (2010) *Rhodococcus equi* infection. **Lancet Infectious Diseases** 10(5): 350-359.

Yapeng, C., Zheng, F., Jianfen, P. and Shijun, Q. (2013) Effects of sulfate residues on the interaction of basic fibroblast growth factor with heparin evaluated by combination of computer modeling and surface plasmon resonance methods. **International Conference on Computer, Networks and Communication Engineering** 297-300

Zink, M. C., Yager, J. A. and Smart, N. L. (1986) *Corynebacterium equi* infections in horses, 1958-1984: A Review of 131 Cases. **Canadian Veterinary Journal** 27(5): 213-217.

Zink, M. C., Yager, J. A., Prescott, J. F. and Fernando, M. A. (1987) Electron microscopic investigation of intracellular events after ingestion of *Rhodococcus equi* by foal alveolar macrophages. **Veterinary Microbiology** 14(3): 295-305.

APPENDICES

Appendix A

MEDIA RECIPES

Luria-Bertani (LB) Broth

10.0 g Tryptone (Oxoid)

5.0 g Yeast extract (Oxoid)

10.0 g NaCl (Sigma)

Dissolve in 1 litre distilled water and autoclave

Brain Heart Infusion (BHI) Broth

47 g BHI broth powder (Oxoid)

Dissolve in 1 litre distilled water and autoclave

Brain Heart Infusion containing sorbitol (BHIS)

BHI supplemented with 2 % sorbitol (Fisher)

Terrific Broth

12 g Bacto Tryptone

24 g Bacto Yeast Extract

900 mL Milli-Q water and autoclave. Cool to about 60°C

Add 100 mL of 0.17 M KH_2PO_4, 0.72 M K_2HPO_4 solution (separately autoclaved).

Solid media

Solid media used were LB agar and BHI agar.

For LB agar, the initial broth was supplemented with 2 % Agar (Bacteriological No. 1) and autoclaved.

For BHI agar, 4.72 g of BHI agar powder (Oxoid) was dissolved in 100 mL of distilled water and autoclaved.

For both media, the autoclaved medium was cooled to about 50°C and aseptically poured into Petri dishes and allowed to set at room temperature.

Special Medium (for growing electro-competent cells)

1 g Tryptone

0.5 g Yeast Extract

1 g NaCl (Sigma)

Dissolve in 80 mL distilled water and autoclave (LB base)

20 mL of distilled water (containing glycine (10 % w/v), Tween-80 (0.5 % v/v)) to LB base (cooled to ~ 60°C) through a 0.22 μm filter.

Selective media

Dissolve recipes for liquid and solid media in water (as above). After autoclaving, each medium was allowed to cool to about 55°C and filter-sterilised antibiotic stock solution was added to the desired concentration and gently shaken. Antibiotics used include ampicillin (100 μg/mL final concentration), and kanamycin (25 μg/mL final concentration)

Appendix B

Table IIA Relevant features of plasmid vectors

Plasmid	Size	Relevant features
pET 23a	3666 bp	Amp^R, T7 tag, T7 term, pRB 322 *ori*
pET 28a	5369 bp	Kan^R, T7 tag, T7 lac, pRB 322 *ori*
pEKEx2	8161 bp	Kan^R, P_{tac}, *lacI*^{q}, pBL1 *ori*V_{Cg}, pUC18 *ori*V_{Ec}

Appendix C

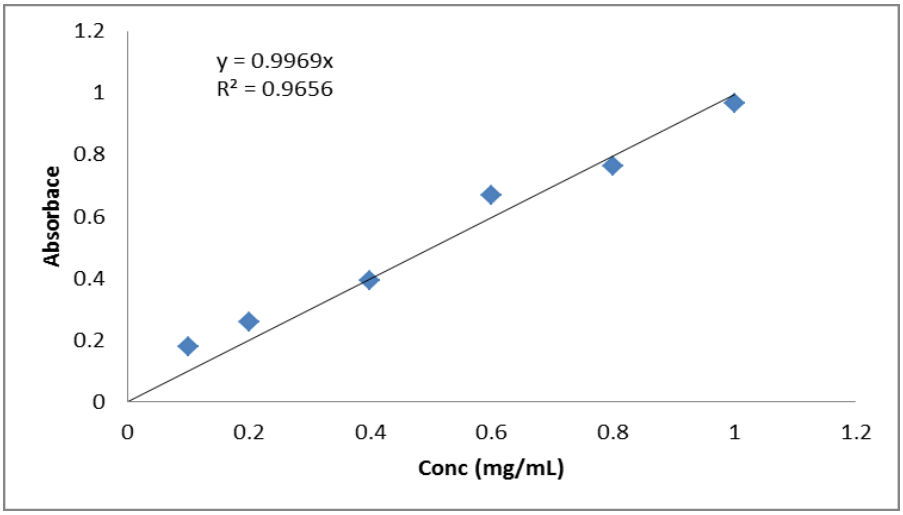


Figure C1: Standard curve for protein quantitation via the Bradford method

Bovine serum albumin was used as the standard

Appendix D

List of buffers and solutions

GTE-RNase buffer (for *C. glutamicum* genomic DNA extraction)

200 µL RNase A

20 mL GTE buffer (50 mM glucose, 25 mM Tris pH 8, 10 mM EDTA pH 8)

DNA Loading buffer (6x)

0.025 g Bromophenol blue

0.025 g Xylene cyanol FF

3 mL glycerol

7 mL deionised water.

50 x Gel buffer or Tris-Acetic acid-EDTA buffer (TAE)

242 g Tris.HCl, pH 8

57.1 mL Glacial acetic acid

100 mL of 0.5 M EDTA

Made to 1 L with deionised water

SDS-PAGE 4x Resolving gel buffer

45.41 g of Tris-HCl pH 8.8

1 g Sodium dodecyl sulphate (SDS)

Made to 250 mL with deionised water

SDS-PAGE 4x Stacking gel buffer

6.06 g of Tris.HCl, pH 6.8

0.4 g SDS

Made to 100 mL with deionised water

SDS-PAGE Sample loading buffer (5x)

0.757 g Tris.HCl, pH 6.8

2 g SDS

10 mL glycerol,

0.15 g Dithiothreitol

0.025 g Bromophenol blue

Made to 100 mL with deionised water

SDS-PAGE running buffer (1x)

3.03 g of Tris base

1 g SDS

14.4 g Glycine

Made to 1 L with deionised water

Native-PAGE 4x Resolving gel buffer

45.41 g of Tris-HCl pH 8.8

Made to 250 mL with deionised water

Native-PAGE 4x Stacking gel buffer

6.06 g of Tris.HCl, pH 6.8

Made to 100 mL with deionised water

Native-PAGE Sample loading buffer (5x)

0.757 g Tris.HCl, pH 6.8

10 mL glycerol,

0.025 g Bromophenol blue

Made to 100 mL with deionised water

Urea-PAGE Sample loading buffer

2.42 g Tris.HCl, pH 6.8

75 mL glycerol

3 g Bromophenol blue

Made to 100 mL with deionised water

Native-PAGE running buffer (1x)

3.03 g of Tris base

14.4 g Glycine

Made to 1 L with deionised water

Coomassie blue stain

0.25 g of Coomassie brilliant blue R-250

45% mL methanol

45% mL Distilled H_2O

10 mL Acetic acid

Destain

100 mL Methanol

100 mL Acetic acid

800 mL Distilled water.

10 x Phosphate buffered saline (PBS)

80 g NaCl

2 g KCl

14.4 g Na_2HPO_4

2.4 g KH_2PO_4

pH 7.4 using NaOH, made to 1 L using deionised water and autoclaved.

PBS was diluted 1 in 10 times before use.

Phosphate buffered saline containing Tween-20 (PBST):

500 mL PBS (1x)

250 µL Tween-20

Transfer buffer:

3.03 g of Tris

14.4 g Glycine

200 mL Absolute Methanol

Made to 1 litre with deionised water

Blocking solution: This is PBST containing 2 % of dried skimmed milk.

Tris buffered saline (TBS buffer)

0.61 g of Tris.HCl, pH 7.4

4.39 g NaCl

Made to 500 mL with deionised water and autoclaved

IMAC equilibration buffer

2.42 g Tris.HCl pH 7.4

2.04 g Imidazole

Made to 1 Litre with deionised water.

IMAC elution buffer (0.02 M Na_2HPO_4, 0.5 M Imidazole):

2.42 g Tris.HCl pH 7.4

34.04 g or 68.08 g Imidazole

Made to 1 Litre with deionised water.

Anion exchange equilibration buffer

2.42 g of Tris. HCl, pH 8.02

Made to 1 Litre with deionised water

Anion exchange elution buffer

2.42 g of Tris.HCl, pH 8.02

58.4 g NaCl

Made to 1 Litre with deionised water

Electroporation buffers (Tris.Glycerol Buffer)

0.12 g Tris.HCl, pH 7.5

100 mL glycerol

Made to 1 Litre with distilled water and autoclaved

Molecular weight markers for SDS-PAGE

α-Lactalbumin from bovine milk	14.2 kDa

Trypsin Inhibitor from soybean	20 kDa
Carbonic Anhydrase from bovine erythrocytes	29 kDa
Glyceraldehyde-3-phosphate Dehydrogenase from rabbit muscle	36 kDa
Ovalbumin from chicken egg	45 kDa

Appendix E

OLIGO DESIGN AND SIGNALP OF *R. EQUI* VAPS

VapA

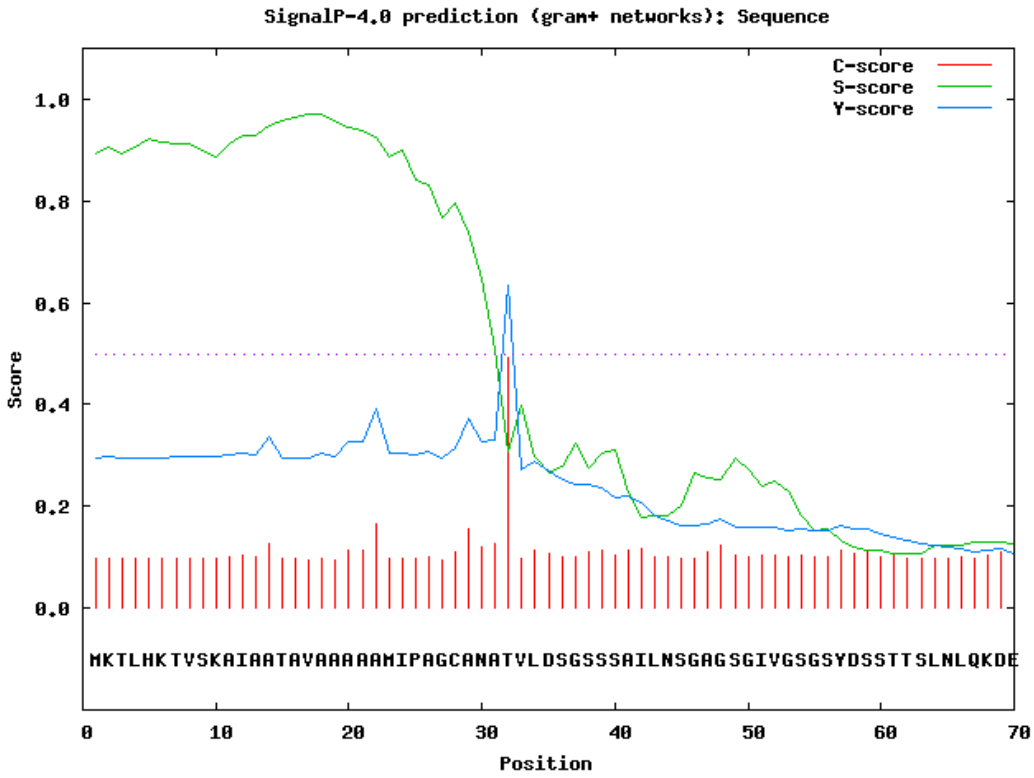


Figure E1 Prediction of signal peptidase cleavage position of VapA using the SignaP 4.0 server.

Figure shows maximum cleavage probability between Alanine (31) and Threonine (32)

VapA

MKTLHKTVSKAIAATAVAAAAAMIPAGCANA**T**VLDSGSSSAILNSGAGSGIVGSGSYDSSTT
SLNLQKDEPNGRASDTAGQEQQYDVHGDVISAVVYQRFHVFGPEGKVFDGDAGGLTLPGAGA
FWGTLFTNDLQRLYKDTVSFQYNAVGPYLNINFFDSSGSFLGHIQSGGVSTVVGVGGGSGSW
HNA

"T" is the predicted N-terminal amino acid of the fully processed protein

VapA

ATGAAGACTCTTCACAAGACGGTTTCTAAGGCGATCGCAGCCACAGCCGTAGCTGCGGCTGC
GGCTATGATTCCCGCCGGCTGCGCTAATGCG**ACC**GTTCTTGATTCCGGTAGCAGCAGTGCG
ATTCTCAATAGTGGGGCAGGCAGTGGCATTGTCGGTTCTGGGAGCTATGACAGCTCGACGAC
TTCGTTAAACCTTCAGAAAGACGAACCGAACGGTCGAGCAAGCGATACCGCCGGGCAAGAGC
AGCAGTACGACGTTCACGGAGACGTCATCAGCGCGGTCGTCTACCAGAGGTTTCACGTATTC
GGGCCAGAAGGAAAGGTCTTCGATGGCGATGCAGGGGGACTCACGCTTCCTGGGGCCGGCGC
GTTCTGGGGGACTCTCTTCACAAATGACCTTCAGCGTCTCTACAAAGACACCGTCTCGTTCC
AGTACAACGCCGTGGGGCCATACCTGAACATCAACTTCTTCGATAGCTCAGGTAGCTTCCTC
GGCCATATCCAGTCCGGTGGAGTTAGTACTGTGGTGGGCGTCGGCGGCGGCTCTGGTAGCTG
GCACAACGCCTAG

"ACC" is the codon for threonine at the predicted N-terminal amino acid of the fully processed protein

VapA without the signal sequence gene

ACCGTTCTTGATTCCGGTAGCAGCAGTGCGATTCTCAATAGTGGGGCAGGCAGTGGCATTGT
CGGTTCTGGGAGCTATGACAGCTCGACGACTTCGTTAAACCTTCAGAAAGACGAACCGAACG
GTCGAGCAAGCGATACCGCCGGGCAAGAGCAGCAGTACGACGTTCACGGAGACGTCATCAGC
GCGGTCGTCTACCAGAGGTTTCACGTATTCGGGCCAGAAGGAAAGGTCTTCGATGGCGATGC
AGGGGGACTCACGCTTCCTGGGGCCGGCGCGTTCTGGGGGACTCTCTTCACAAATGACCTTC
AGCGTCTCTACAAAGACACCGTCTCGTTCCAGTACAACGCCGTGGGGCCATACCTGAACATC
AACTTCTTCGATAGCTCAGGTAGCTTCCTCGGCCATATCCAGTCCGGTGGAGTTAGTACTGT
GGTGGGCGTCGGCGGCGGCTCTGGTAGCTGGCACAACGCCTAG

pET23vapA up	GATCGATC**CATATG**ACCGTTCTTGATTCCGGTAGCAGCAGTGCG
pET23vapA lo	TATAAATA**GCGGCCGC**GGCGTTGTGCCAGCTACCAGAGCCG

Nde1…. **CATATG**

Not1…. **GCGGCCGC**

VapC

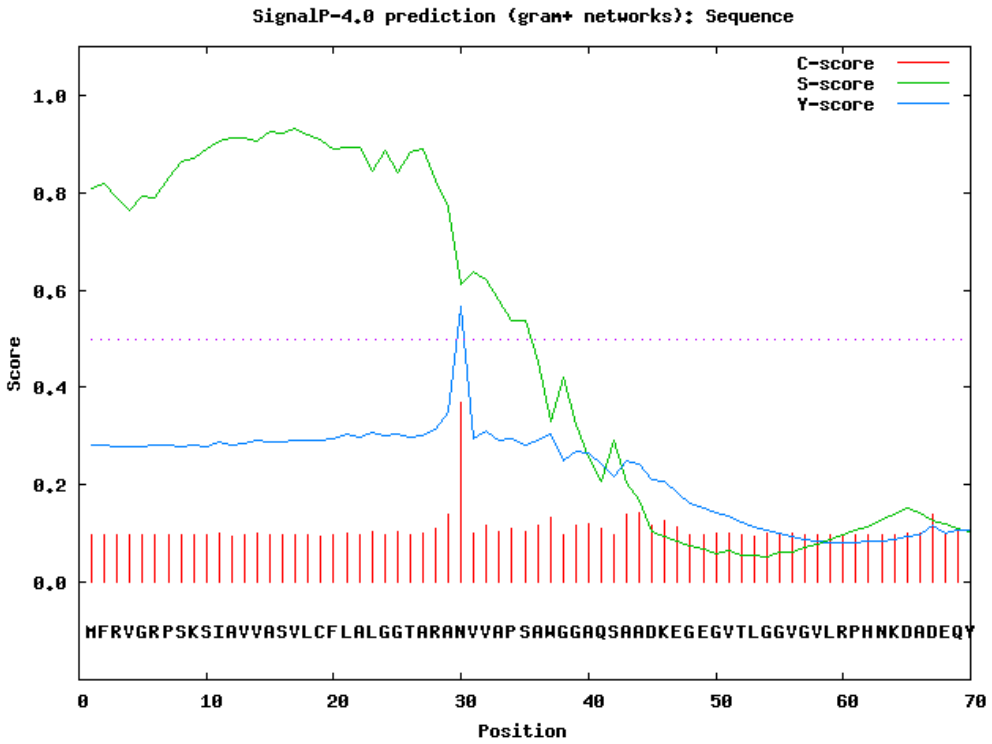


Figure E2 Prediction of signal peptidase cleavage position of VapC using the SignaP 4.0 server.

Figure shows maximum cleavage probability between Alanine (29) and Asparagine (30)

VapC
MFRVGRPSKSIAVVASVLCFLALGGTARA**N**VVAPSAWGGAQSAADKEGEGVTLGGVGVLRPH
NKDADEQYTVHGVVVSALFYNHLRISVDGGMTFDGDGGGLSTPGGGALWGTLTTSDLQQLYD
ETASFECNAVGPYLNINFYDSYGRILASVQAGGVSTMIGIGGGNGRWHLV

"N" (30) is the predicted N-terminal amino acid of the fully processed protein

VapC

ATGTTCCGCGTGGGTCGTCCATCCAAATCGATAGCCGTTGTGGCGAGTGTCCTATGCTTCTT
GGCCTTGGGAGGCACTGCACGGGCC**AAT**GTAGTCGCTCCGTCGGCGTGGGGAGGGGCGCAA
TCCGCTGCTGACAAAGAGGGCGAAGGGGTCACGCTTGGTGGCGTCGGAGTTCTACGGCCGCA
CAATAAAGATGCCGACGAACAGTATACGGTTCATGGGGTAGTTGTCAGTGCGTTGTTTTATA
ATCATCTGCGGATCTCTGTCGATGGGGGAATGACCTTCGATGGAGACGGAGGTGGGCTATCC
ACTCCGGGAGGAGGTGCGCTCTGGGGAACTCTTACAACTAGCGATCTTCAGCAACTATACGA
CGAAACTGCGTCCTTCGAATGTAATGCGGTTGGCCCGTATTTGAACATCAACTTTTACGACA
GTTACGGTCGCATTCTTGCGAGTGTTCAGGCGGGCGGTGTCAGCACAATGATTGGTATAGGC
GGTGGGAATGGGCGATGGCATTTGGTGTGA

"AAT" is the codon for asparagine at the predicted N-terminal amino acid of the fully processed protein

VapC without signal sequence gene
AATGTAGTCGCTCCGTCGGCGTGGGGAGGGGCGCAATCCGCTGCTGACAAAGAGGGCGAAGG
GGTCACGCTTGGTGGCGTCGGAGTTCTACGGCCGCACAATAAAGATGCCGACGAACAGTATA
CGGTTCATGGGGTAGTTGTCAGTGCGTTGTTTTATAATCATCTGCGGATCTCTGTCGATGGG
GGAATGACCTTCGATGGAGACGGAGGTGGGCTATCCACTCCGGGAGGAGGTGCGCTCTGGGG
AACTCTTACAACTAGCGATCTTCAGCAACTATACGACGAAACTGCGTCCTTCGAATGTAATG
CGGTTGGCCCGTATTTGAACATCAACTTTTACGACAGTTACGGTCGCATTCTTGCGAGTGTT
CAGGCGGGCGGTGTCAGCACAATGATTGGTATAGGCGGTGGGAATGGGCGATGGCATTTGGT
GTGA

For cloning into pET23
Forward primer 5'-GATCGATC**CATATG**AATGTAGTCGCTCCGTCGGCGTGG -3'
Reverse primer 5'-AATTTATA**GCGGCCGC**CACCAAATGCCATCGCCCATTCC -3'

Nde1…. **CATATG**

Not1…. **GCGGCCGC**

VapD

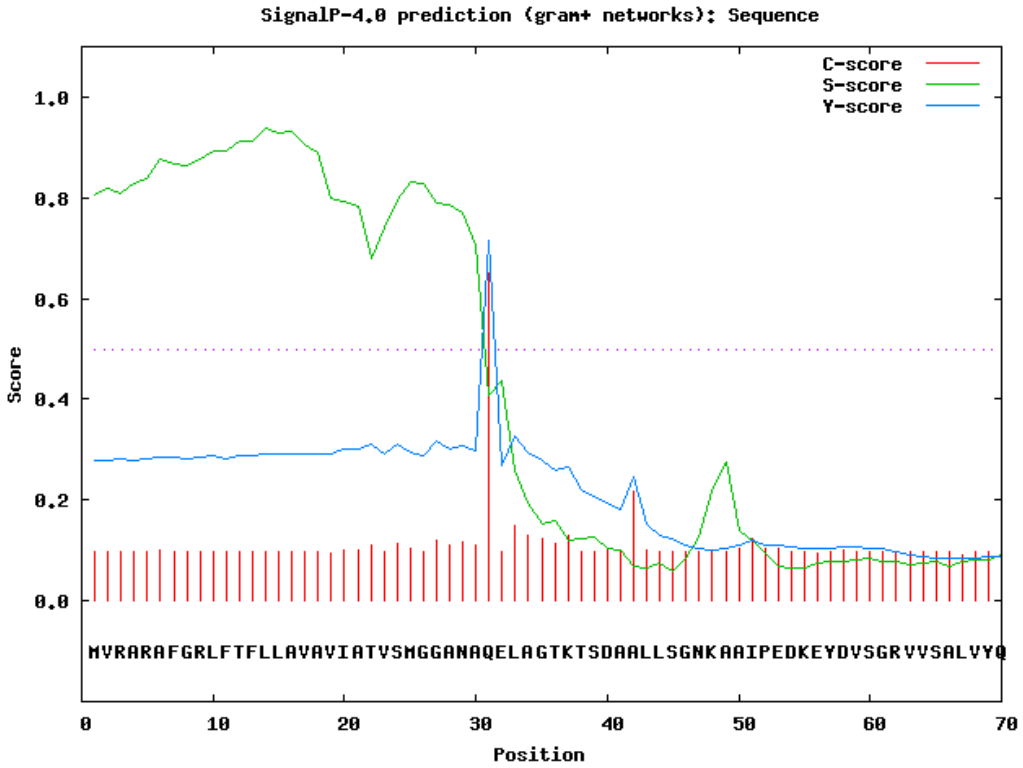


Figure E3 Prediction of signal peptidase cleavage position of VapD using the SignaP 4.0 server.

Figure shows maximum cleavage probability between Alanine (30) and Glutamine (31)

VapD
MVRARAFGRLFTFLLAVAVIATVSMGGANA**Q**ELAGTKTSDAALLSGNKAAIPEDKEYDVSGR
VVSALVYQYFIVTVDDAEDKKGKTFQGDAGGVTIPGVDFFWGTLHTPDLEKLYSDTVSFQYN
AAATFLNINFFDSKGERLGYVLAGAAGTVSGIGGGTGGWE

"Q" (31) is the predicted N-terminal amino acid of the fully processed protein

VapD

ATGGTCCGTGCACGAGCCTTTGGGCGTTTATTCACTTTCTTGCTCGCGGTGGCTGTGATCGC
AACGGTCTCAATGGGAGGGGCCAACGCT**CAG**GAGCTAGCTGGCACCAAGACGTCGGATGCG
GCACTTCTCTCCGGAAACAAGGCTGCCATTCCTGAAGATAAAGAGTATGACGTATCGGGTAG
GGTTGTCAGTGCTCTAGTCTACCAATATTTCATTGTAACCGTCGATGATGCTGAGGATAAGA
AGGGCAAGACGTTCCAGGGGGACGCTGGGGGTGTGACCATTCCGGGAGTCGACTTCTTCTGG

GGTACTTTGCACACGCCGGACCTTGAAAAGCTGTATTCCGATACAGTGTCGTTTCAATACAA
TGCGGCCGCAACATTCTTGAATATCAACTTTTTCGACAGCAAGGGCGAACGACTTGGTTACG
TTCTTGCTGGTGCTGCTGGGACAGTATCGGGGATCGGTGGTGGCACTGGCGGGTGGGAGTAG

"CAG" is the codon for glutamine at the predicted N-terminal amino acid of the fully processed protein

VapD without signal sequence

CAGGAGCTAGCTGGCACCAAGACGTCGGATGCGGCACTTCTCTCCGGAAACAAGGCTGCCAT
TCCTGAAGATAAAGAGTATGACGTATCGGGTAGGGTTGTCAGTGCTCTAGTCTACCAATATT
TCATTGTAACCGTCGATGATGCTGAGGATAAGAAGGGCAAGACGTTCCAGGGGGACGCTGGG
GGTGTGACCATTCCGGGAGTCGACTTCTTCTGGGGTACTTTGCACACGCCGGACCTTGAAAA
GCTGTATTCCGATACAGTGTCGTTTCAATACAATGCGGCCGCAACATTCTTGAATATCAACT
TTTTCGACAGCAAGGGCGAACGACTTGGTTACGTTCTTGCTGGTGCTGCTGGGACAGTATCG
GGGATCGGTGGTGGCACTGGCGGGTGGGAGTAG

For cloning into pET28

Forward Primer 5'-GTAGGAAC**CATATG**CAGGAGCTAGCTGGCACCAAGAC-3'

Reverse primer 5'-GTAGGAAC**GAATTC**CTACTCCCACCCGCCAGTG-3'

CATATG……**NdeI site**

GAATTC……**EcoRI site**

VapE

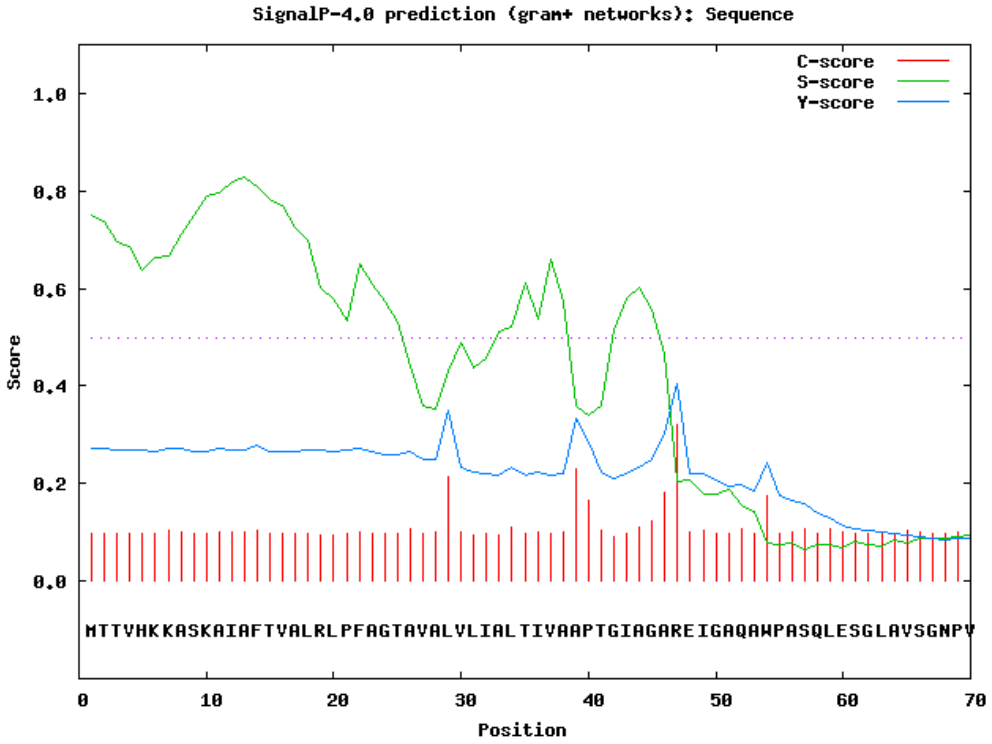


Figure E4 Prediction of signal peptidase cleavage position of VapE using the SignaP 4.0 server.

Figure shows maximum cleavage probability between Alanine (46) and Arginine (47)

VapE

MTTVHKKASKAIAFTVALRLPFAGTAVALVLIALTIVAAPTGIAGA**R**EIGAQAWPASQLES
GLAVSGNPVGVHDVRMAVHDDSTHTREFKEDDSEKQYPVHGFASSFIFYQTVSIIIDDDGRG
GPGKTFEGEAGGITTPGAAGYAGVLFTSDLERLYRETVSFEYNAVGPYLNINLFAGDGGLLG
HVQSGAISSLVGIGGGTGAWR

"R" (47) is the predicted N-terminal amino acid of the fully processed protein

VapE

ATGACGACCGTTCACAAGAAAGCCTCTAAGGCAATCGCCTTCACAGTCGCATTGCGGCTCCC
TTTCGCGGGAACCGCAGTAGCATTGGTTCTAATCGCACTGACTATAGTAGCTGCTCCTACAG
GCATAGCCGGCGCG**CGG**GAAATTGGCGCCCAGGCCTGGCCAGCTTCGCAATTGGAAAGCGG
CCTTGCTGTGTCGGGAAATCCTGTGGGCGTACATGACGTCCGTATGGCGGTGCATGACGATT
CCACGCACACGCGGGAGTTTAAGGAAGATGACTCAGAGAAGCAGTATCCGGTACATGGCTTT
GCTTCGAGCTTTATCTTCTATCAAACCGTATCAATAATAATTGATGACGACGGGCGCGGAGG
GCCGGGTAAGACGTTCGAAGGAGAAGCTGGCGGTATTACCACGCCCGGCGCCGCTGGTTACG
CCGGTGTTCTATTCACGAGCGATCTCGAGCGGCTCTACCGCGAGACGGTTTCGTTTGAGTAC
AACGCTGTCGGTCCGTACTTGAACATCAATCTTTTCGCCGGAGACGGAGGTTTGCTGGGGCA
CGTCCAATCTGGCGCTATCAGTAGTTTGGTGGGCATCGGAGGAGGCACTGGCGCCTGGCGAT
AG

"**CGG"** is the codon for arginine at the predicted N-terminal amino acid of the fully processed protein

VapE without signal sequence

CGGGAAATTGGCGCCCAGGCCTGGCCAGCTTCGCAATTGGAAAGCGGCCTTGCTGTGTCGGG
AAATCCTGTGGGCGTACATGACGTCCGTATGGCGGTGCATGACGATTCCACGCACACGCGGG
AGTTTAAGGAAGATGACTCAGAGAAGCAGTATCCGGTACATGGCTTTGCTTCGAGCTTTATC
TTCTATCAAACCGTATCAATAATAATTGATGACGACGGGCGCGGAGGGCCGGGTAAGACGTT
CGAAGGAGAAGCTGGCGGTATTACCACGCCCGGCGCCGCTGGTTACGCCGGTGTTCTATTCA
CGAGCGATCTCGAGCGGCTCTACCGCGAGACGGTTTCGTTTGAGTACAACGCTGTCGGTCCG
TACTTGAACATCAATCTTTTCGCCGGAGACGGAGGTTTGCTGGGGCACGTCCAATCTGGCGC
TATCAGTAGTTTGGTGGGCATCGGAGGAGGCACTGGCGCCTGGCGATAG

For cloning into pET28

Forward primer 5'-GTACGATC**CATATG**CGGGAAATTGGCGCCCAG-3'

Reverse primer 5'-GTACGATC**GAATTC**CTATCGCCAGGCGCCAGTG-3'

Nde1…. **CATATG**

EcoR1… **GAATTC**

VapG

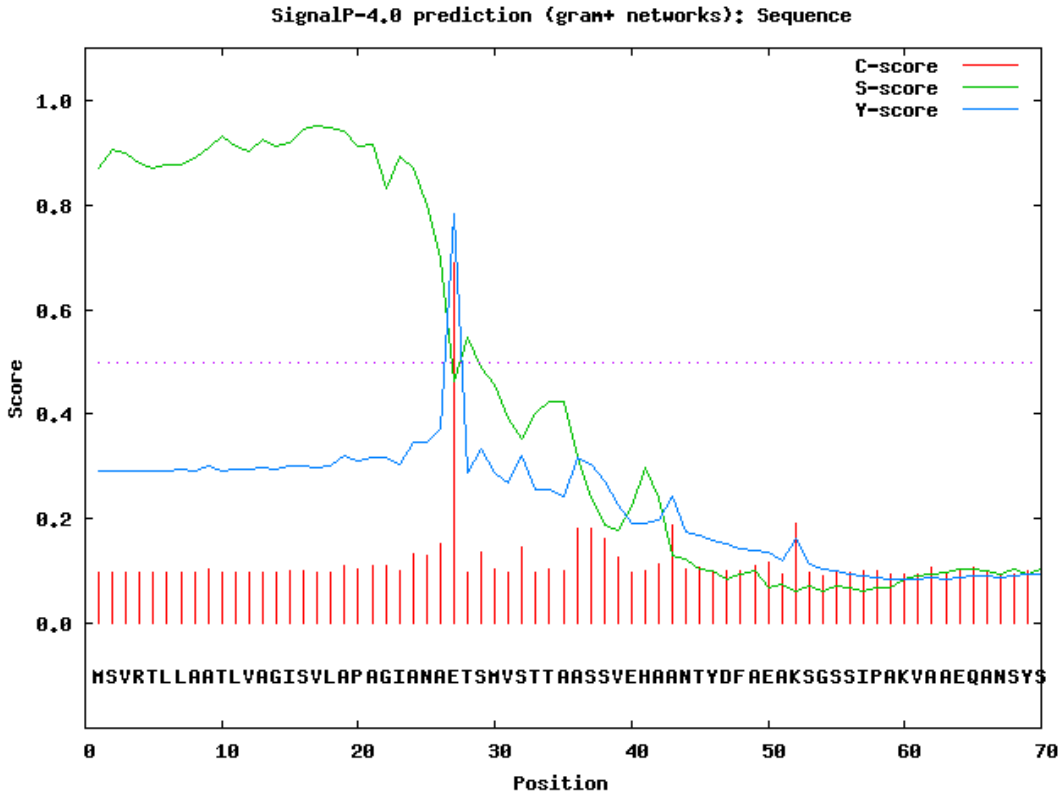


Figure E5: Prediction of signal peptidase cleavage position of VapG using the SignaP 4.0 server.

Figure shows maximum cleavage probability between Alanine (26) and glutamic acid (27)

VapG

MSVRTLLAATLVAGISVLAPAGIANA**E**TSMVSTTAASSVEHAANTYDFAEAKSGSSIPAKVA
AEQANSYSVHGLVTSLAVYQHFSLTVEGGGKTFTGDSGGISIPGVAVLEGTLFTEDLQHLYS
DTVSFEYNAVGPYLNINFFDSHGTLLGHVQSGSIGTVSGIGGGTGGWQ

"E" (27) is the predicted N-terminal amino acid of the fully processed protein

VapG

GTGAGTGTTCGGACCCTTTTGGCGGCAACGCTCGTTGCTGGAATATCAGTCTTGGCACCGGC
CGGCATTGCGAACGCG**GAA**ACTTCAATGGTATCCACTACAGCAGCATCGAGTGTCGAGCAC
GCTGCAAACACCTACGACTTTGCAGAGGCGAAGAGCGGGAGCTCTATCCCCGCCAAAGTAGC
CGCAGAGCAGGCAAACAGCTATTCGGTCCACGGGCTTGTCACCAGCCTCGCCGTATACCAGC
ACTTTTCACTGACCGTTGAAGGCGGCGGAAAGACGTTTACTGGTGATTCTGGCGGGATTTCG
ATTCCCGGGGTTGCAGTGCTGGAGGGAACCCTATTCACCGAGGATCTGCAGCATTTGTACAG
CGACACCGTCTCGTTCGAGTACAACGCCGTAGGCCCGTACCTGAACATCAACTTTTTTGACA
GCCATGGCACTCTCCTAGGCCACGTGCAGTCTGGATCCATCGGGACCGTCTCCGGCATCGGT
GGCGGAACCGGAGGGTGGCAATAG

"GAA" is the codon for glutamic acid at the predicted N-terminal amino acid of the fully processed protein

VapG without signal sequence

GAAACTTCAATGGTATCCACTACAGCAGCATCGAGTGTCGAGCACGCTGCAAACACCTACGA
CTTTGCAGAGGCGAAGAGCGGGAGCTCTATCCCCGCCAAAGTAGCCGCAGAGCAGGCAAACA
GCTATTCGGTCCACGGGCTTGTCACCAGCCTCGCCGTATACCAGCACTTTTCACTGACCGTT
GAAGGCGGCGGAAAGACGTTTACTGGTGATTCTGGCGGGATTTCGATTCCCGGGGTTGCAGT
GCTGGAGGGAACCCTATTCACCGAGGATCTGCAGCATTTGTACAGCGACACCGTCTCGTTCG
AGTACAACGCCGTAGGCCCGTACCTGAACATCAACTTTTTTGACAGCCATGGCACTCTCCTA
GGCCACGTGCAGTCTGGATCCATCGGGACCGTCTCCGGCATCGGTGGCGGAACCGGAGGGTG
GCAATAG

For cloning into pET28

Forward primer 5'-GACGATC**CATATG**GAAACTTCAATGGTATCCACTACAGCAGCATCG-3'

Reverse primer 5'-GACGATC**GAATTC**CTATTGCCACCCTCCGGTTCC-3'

Nde1…. **CATATG**

EcoR1… **GAATTC**

VapH

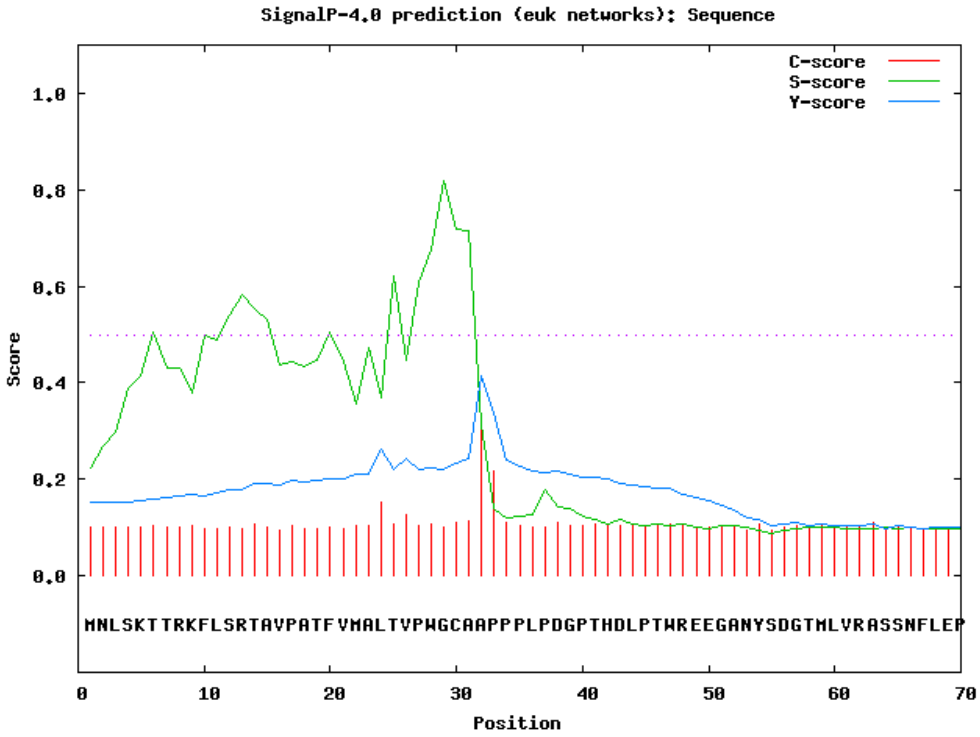


Figure E6 Prediction of signal peptidase cleavage position of VapG using the SignaP 4.0 server.

Figure shows maximum cleavage probability between alanines 31 and 32

VapH
MNLSKTTRKFLSRTAVPATFVMALTVPWGCA**A**PPPLPDGPTHDLPTWREEGANYSDGTMLV
RASSNFLEPSTHSDSGQQQWTVQGVLASALVYQRLKLNVEGGETFEGYAGGLSFPGGAMVWG
TLFTDNIQRLYDRTESFEFNAVGPYLNVNFFDGHSAILGHAQLGGVSSVIGIGGGTGTWIGD
VA

"A" (32) is the predicted N-terminal amino acid of the fully processed protein

VapH

ATGAATCTATCAAAAACAACTCGAAAGTTCCTATCAAGGACAGCTGTGCCTGCAACATTCGT
AATGGCGCTCACTGTCCCGTGGGGATGCGCG**GCC**CCGCCACCATTACCAGATGGTCCCACA
CACGACCTGCCTACTTGGCGCGAAGAAGGGGCAAACTATAGCGACGGTACGATGCTTGTACG
CGCTTCATCCAATTTTCTCGAGCCATCGACTCACAGTGACAGTGGGCAGCAGCAGTGGACAG
TGCAAGGAGTACTGGCCAGCGCGTTGGTGTACCAACGGCTGAAGCTCAATGTTGAGGGAGGC
GAGACGTTTGAGGGTTATGCAGGTGGATTGTCATTTCCCGGCGGGGCGATGGTCTGGGGCAC
GCTTTTCACCGACAACATACAGCGACTATACGATCGGACGGAATCGTTTGAGTTCAACGCGG
TCGGACCCTATCTAAATGTCAACTTCTTCGACGGTCACAGCGCCATACTCGGCCATGCACAA
TTAGGAGGGGTGAGCTCCGTAATCGGTATTGGCGGAGGTACGGGAACCTGGATAGGCGATGT
AGCGTAG

"**GCC**" is the codon for alanine at the predicted N-terminal amino acid of the fully processed protein

VapH without signal sequence

GCCCCGCCACCATTACCAGATGGTCCCACACACGACCTGCCTACTTGGCGCGAAGAAGGGGC
AAACTATAGCGACGGTACGATGCTTGTACGCGCTTCATCCAATTTTCTCGAGCCATCGACTC
ACAGTGACAGTGGGCAGCAGCAGTGGACAGTGCAAGGAGTACTGGCCAGCGCGTTGGTGTAC
CAACGGCTGAAGCTCAATGTTGAGGGAGGCGAGACGTTTGAGGGTTATGCAGGTGGATTGTC
ATTTCCCGGCGGGGCGATGGTCTGGGGCACGCTTTTCACCGACAACATACAGCGACTATACG
ATCGGACGGAATCGTTTGAGTTCAACGCGGTCGGACCCTATCTAAATGTCAACTTCTTCGAC
GGTCACAGCGCCATACTCGGCCATGCACAATTAGGAGGGGTGAGCTCCGTAATCGGTATTGG
CGGAGGTACGGGAACCTGGATAGGCGATGTAGCGTAG

Forward primer 5'-GACGATC**CATATG**GCCCCGCCACCATTACCAGATG-3'

Reverse primer 5'-GACGATC**GAATTC**CTACGCTACATCGCCTATCCAGGTTCCCGTAC-3'

Nde1…. **CATATG**

EcoR1… **GAATTC**

Appendix F

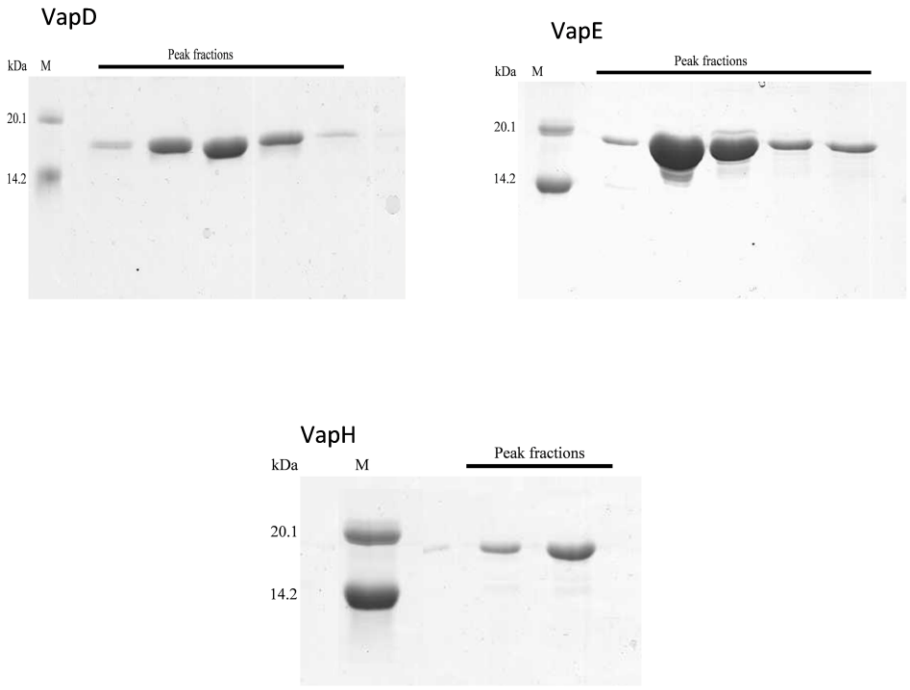


Figure F1 SDS-PAGE of recombinant VapD, VapE and VapH (expressed in *E. coli*) following anion exchange chromatography

Peak fractions (marked) were concentrated and used for further analysis. 'M' indicates molecular weight markers (14.2 kDa, α-Lactalbumin from bovine milk; 20.1 kDa, Trypsin Inhibitor from soybean).

Appendix G

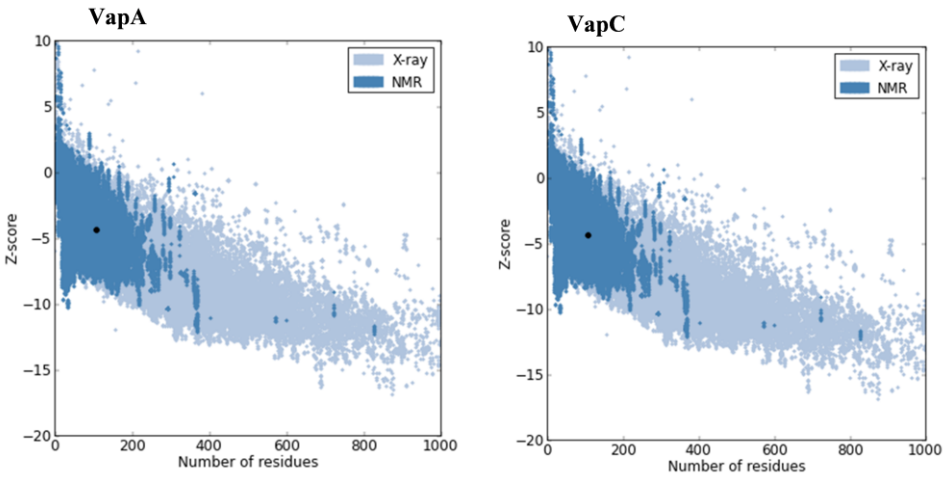


Figure G1: Model quality plots of VapA and VapC showing Z-scores of -4.4 and -4.35 respectively.

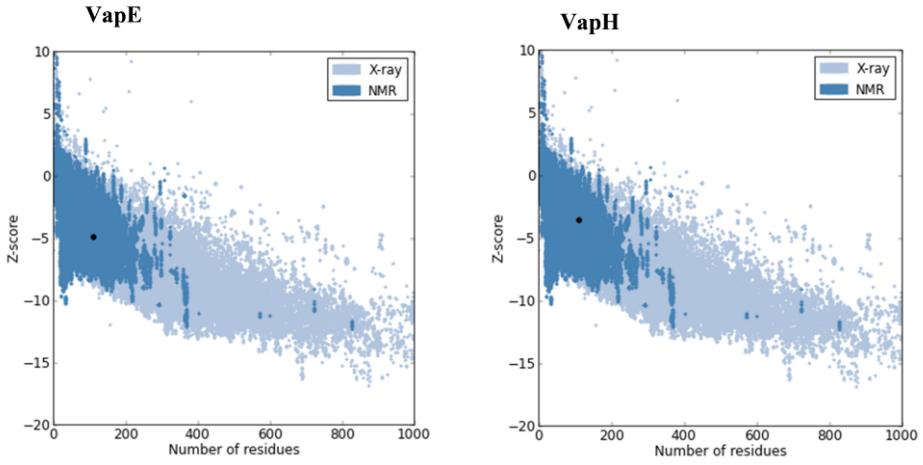


Figure G2: Model quality plots of VapE and VapH showing Z-scores of -4.88 and -3.35 respectively

Appendix H

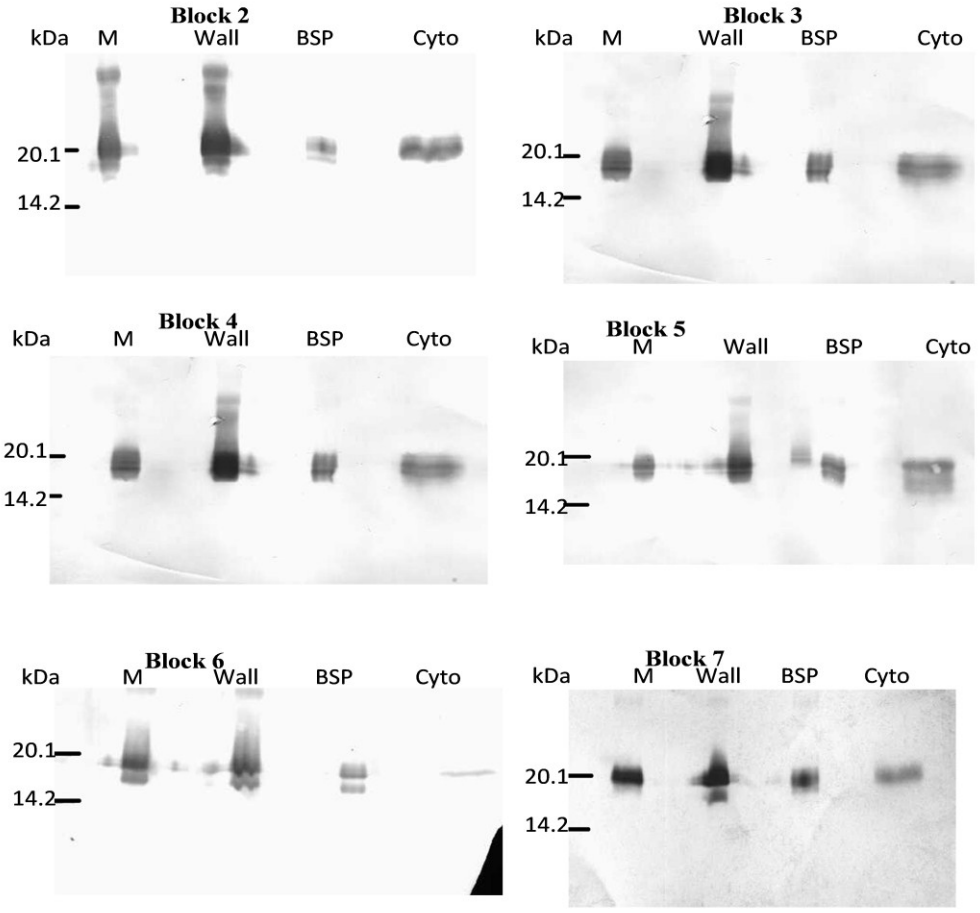


Figure H1 Site-directed mutagenesis did not perturb recombinant VapA localisation in *C. glutamicum*

Fractions were subjected to Western Blot using Mab 10G5 as the primary antibody. Site-directed mutagenesis was performed in blocks as described. Shown are blocks 2 – 7. Fractions are M, plasma membrane; Wall, purified cell wall; BSP, secreted protein; and cyto, cytosol. Standardised loading of 5 µg total protein was used throughout.

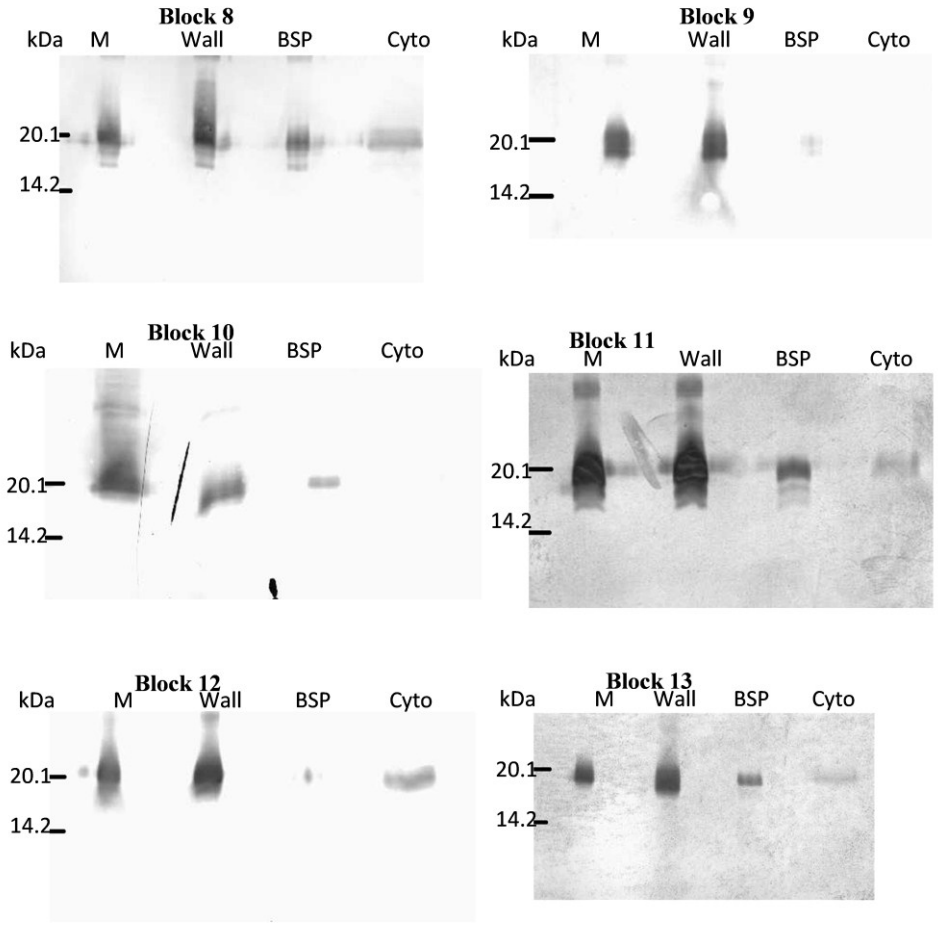


Figure H2 Site-directed mutagenesis did not perturb recombinant VapA localisation in *C. glutamicum*

Fractions were subjected to Western Blot using Mab 10G5 as the primary antibody. Site-directed mutagenesis was performed in blocks as described. Shown are blocks 8 – 13. Fractions are M, plasma membrane; Wall, purified cell wall; BSP, secreted protein; and cyto, cytosol. Standardised loading of 5 µg total protein was used throughout.

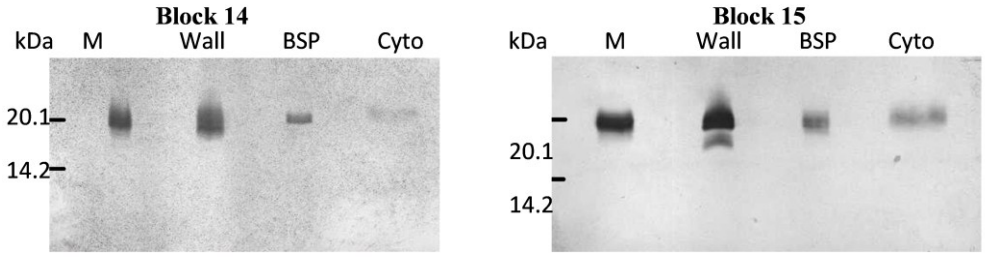


Figure H3 Site-directed mutagenesis did not perturb recombinant VapA localisation in *C. glutamicum*

Fractions were subjected to Western Blot using Mab 10G5 as the primary antibody. Site-directed mutagenesis was performed in blocks as described. Shown are blocks 14 – 15. Fractions are M, plasma membrane; Wall, purified cell wall; BSP, secreted protein; and cyto, cytosol. Standardised loading of 5 µg total protein was used throughout.

Appendix I

Plasmid maps

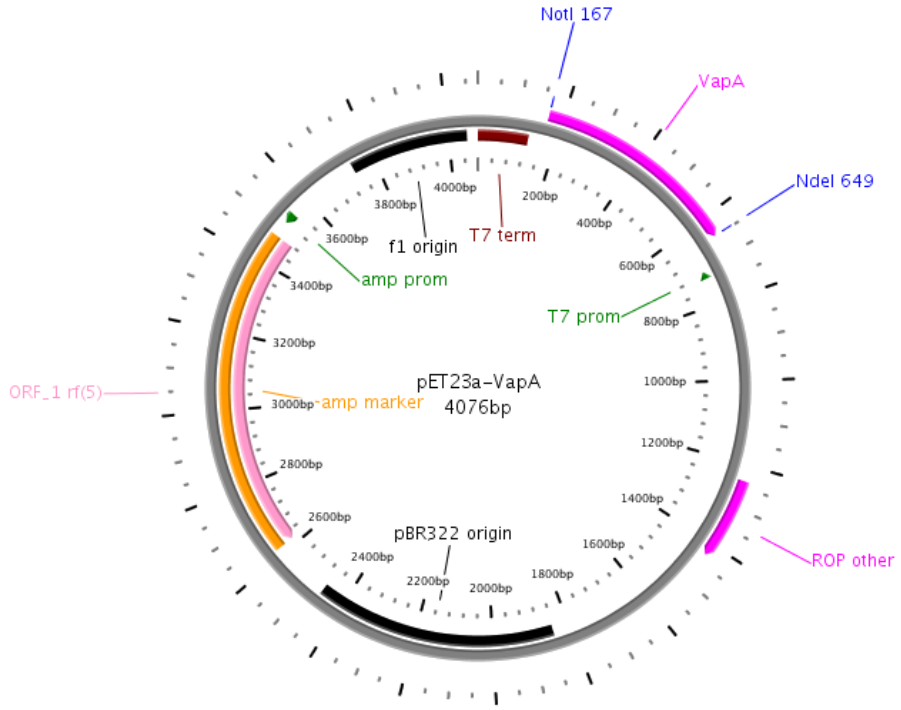


Figure I1: Map of pET23a-*vapA*

Gene (*vapA*) was cloned into the NotI/NdeI cloning sites of pET23a as shown. Map created using PlasMapper (Dong *et al*., 2004).

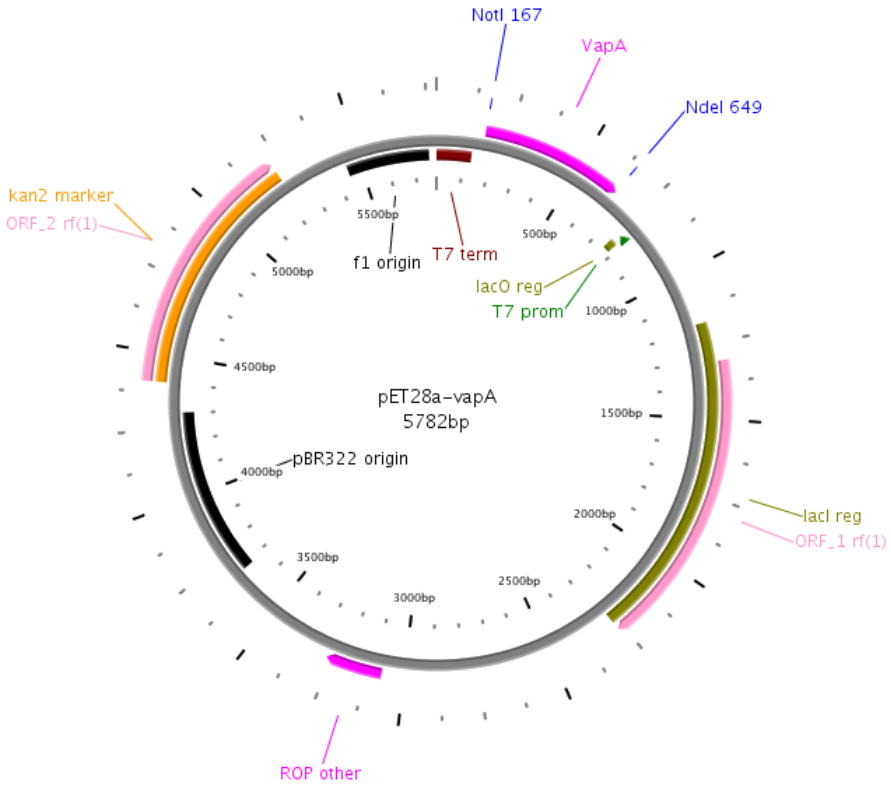


Figure I2: Map of pET28a-*vapA*.

Gene (*vapA*) was cloned into the NotI/NdeI cloning sites of pET28a as shown. Map created using PlasMapper (Dong *et al*., 2004).

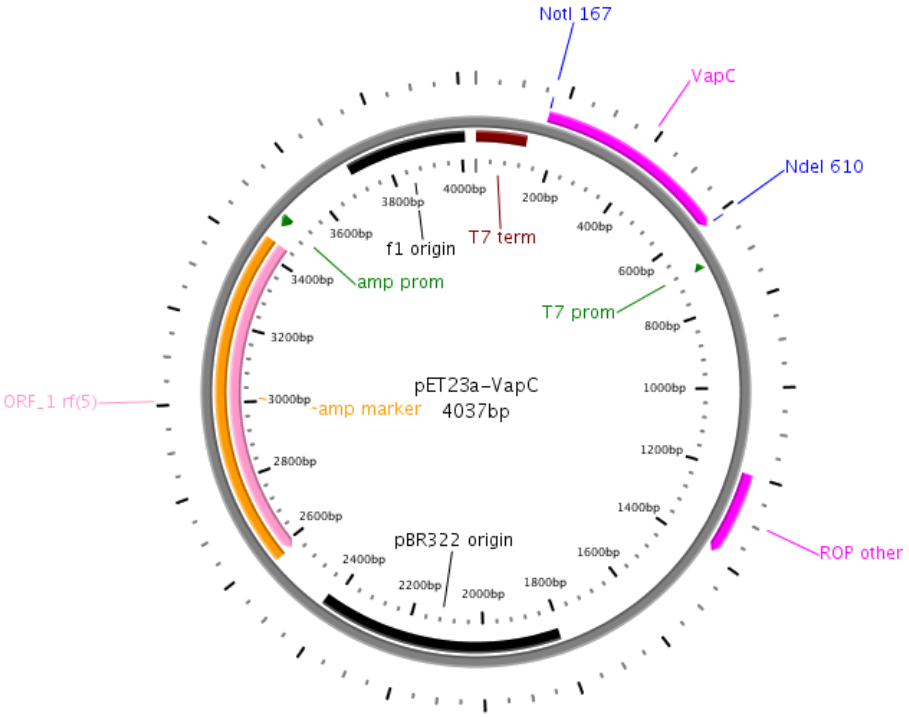


Figure I3: Map of pET23a-*vapC*.

Gene (*vapC*) was cloned into the NotI/NdeI cloning sites of pET23a as shown. Map created using PlasMapper (Dong *et al*., 2004).

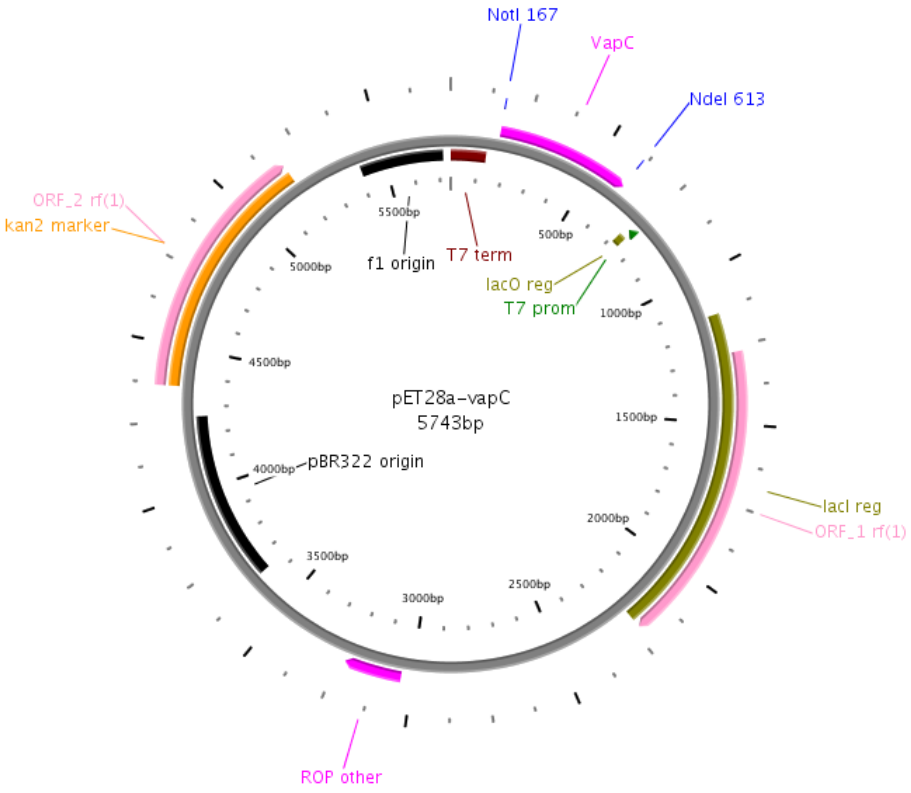


Figure I4: Map of pET28a-*vapC*.

Gene (*vapC*) was cloned into the NotI/NdeI cloning sites of pET28a as shown. Map created using PlasMapper (Dong *et al*., 2004).

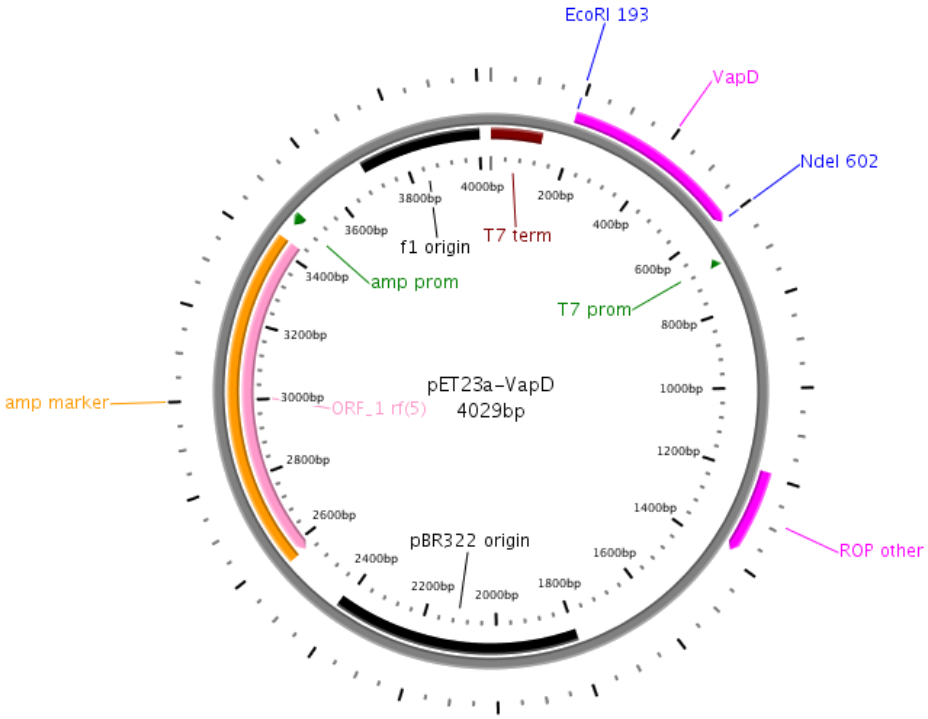


Figure I5: Map of pET23a-*vapD*

Gene (*vapD*) was cloned into the EcoRI/NdeI cloning sites of pET23a as shown. Map created using PlasMapper (Dong *et al*., 2004).

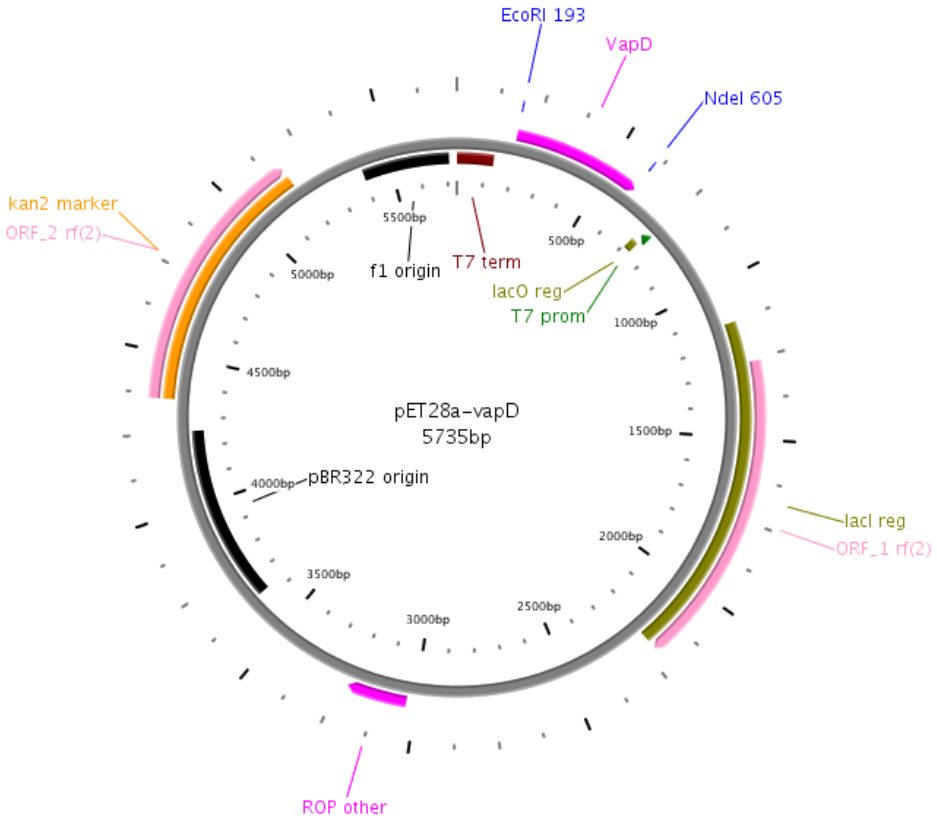


Figure I6: Map of pET28a-*vapD*.

Gene (*vapD*) was cloned into the EcoRI/NdeI cloning sites of pET28a as shown. Map created using PlasMapper (Dong *et al*., 2004).

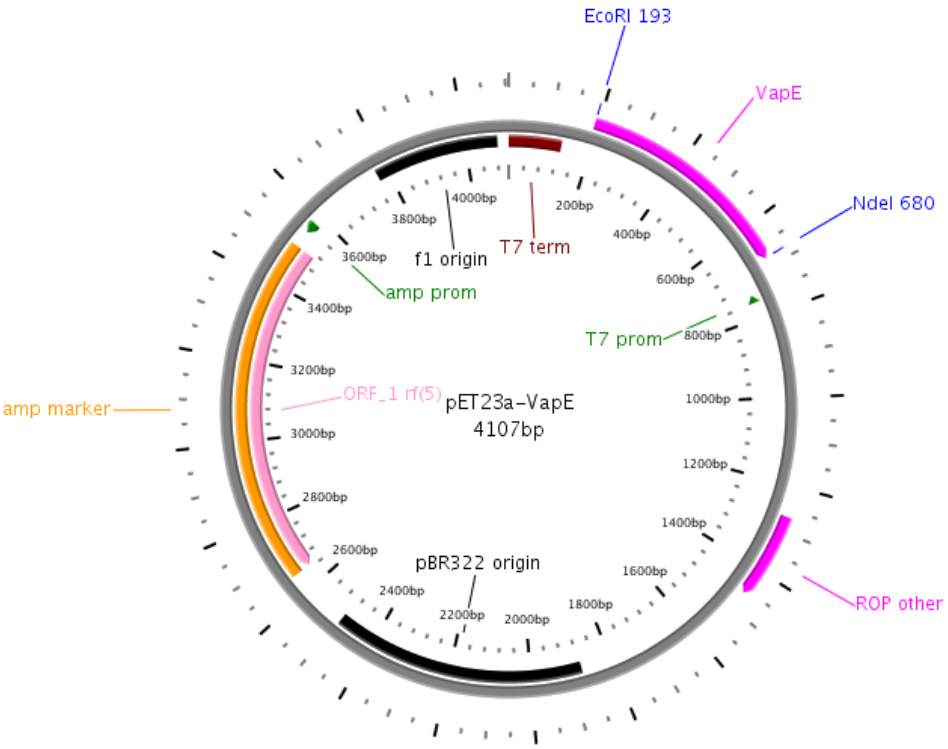


Figure I7: Map of pET23a-*vapE*.

Gene (*vapE*) was cloned into the EcoRI/NdeI cloning sites of pET23a as shown. Map created using PlasMapper (Dong *et al*., 2004).

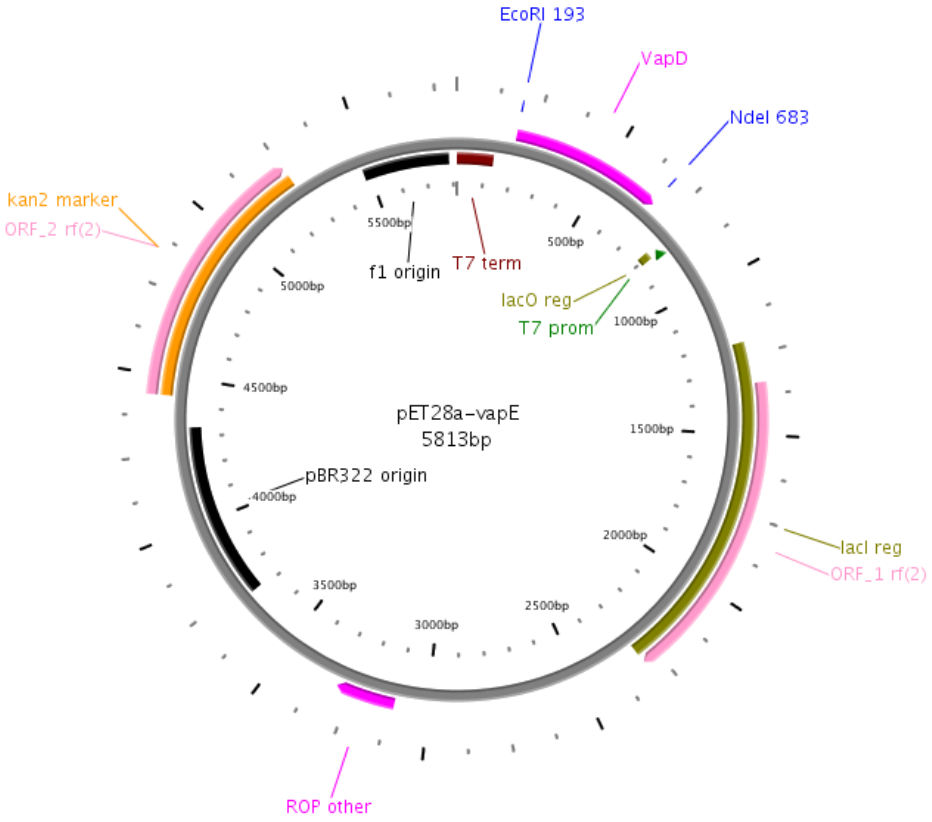


Figure I8: Map of pET28a-*vapE*.

Gene (*vapE*) was cloned into the EcoRI/NdeI cloning sites of pET28a as shown. Map created using PlasMapper (Dong *et al*., 2004).

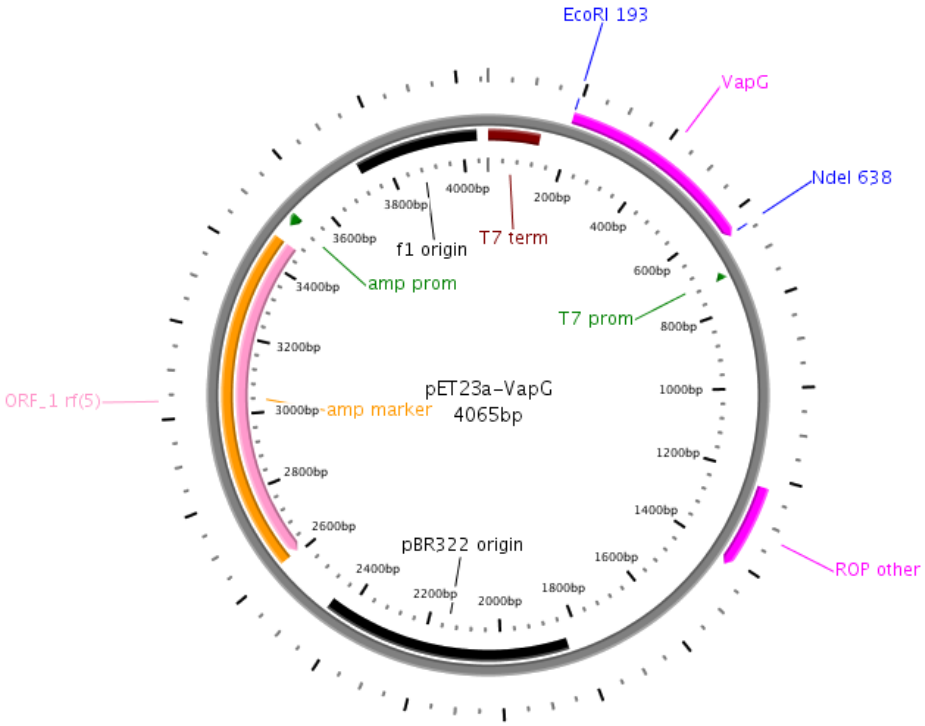


Figure I9: Map of pET23a-*vapG*

Gene (*vapG*) was cloned into the EcoRI/NdeI cloning sites of pET23a as shown. Map created using PlasMapper (Dong *et al*., 2004).

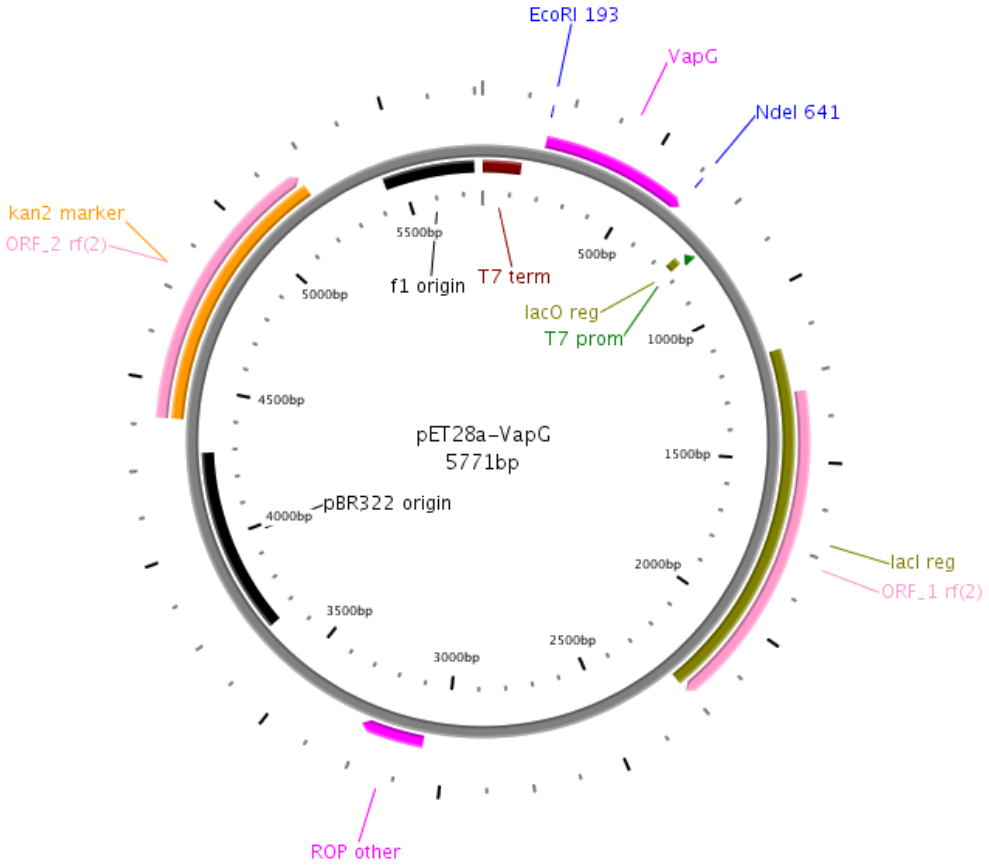


Figure I10: Map of pET28a-*vapG*.

Gene (*vapG*) was cloned into the EcoRI/NdeI cloning sites of pET28a as shown. Map created using PlasMapper (Dong *et al*., 2004).

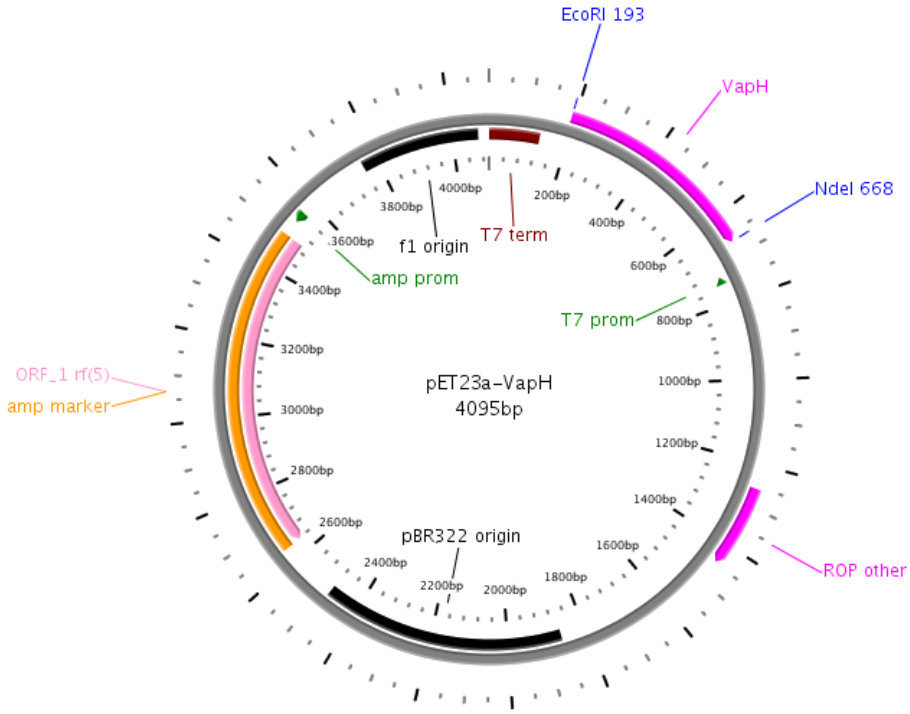


Figure I11 Map of pET23a-*vapH*

Gene (*vapH*) was cloned into the EcoRI/NdeI cloning sites of pET23a as shown. Map created using PlasMapper (Dong *et al*., 2004).

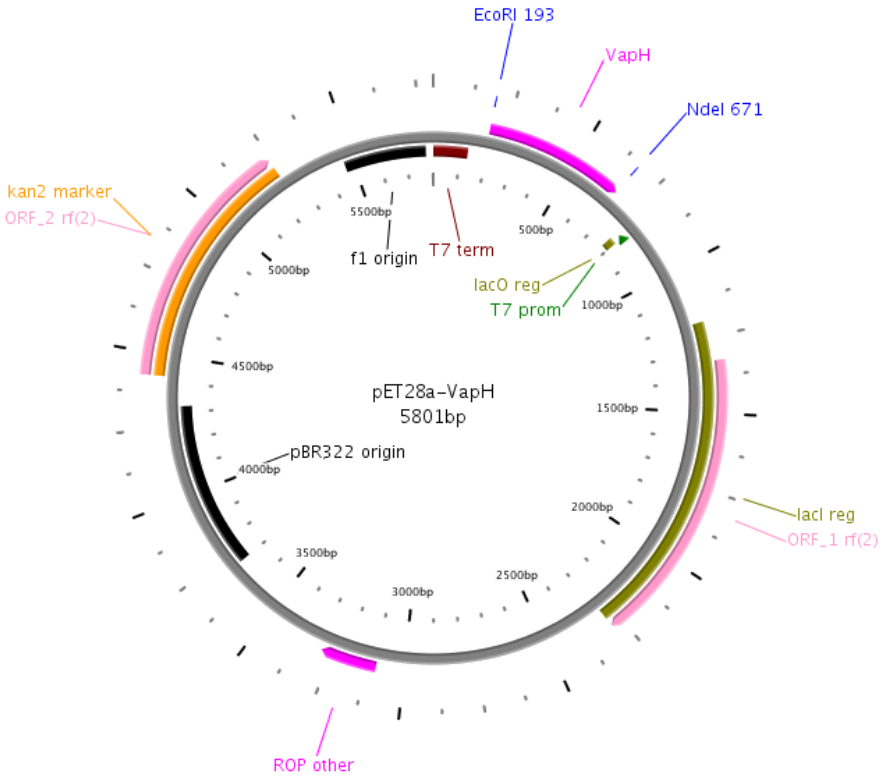


Figure I12: Map of pET28a-*vapH*.

Gene (*vapH*) was cloned into the EcoRI/NdeI cloning sites of pET28a as shown. Map created using PlasMapper (Dong *et al.*, 2004).

Appendix J

Crystallisation Liquors

Table J1: PEG/ION screen components

Tube No.	Salt	Precipitant
1	0.2 M Sodium fluoride	20% w/v PEG 3,350
2	0.2 M Potassium fluoride	20% w/v PEG 3,350
3	0.2 M Ammonium fluoride	20% w/v PEG 3,350
4	0.2 M Lithium chloride	20% w/v PEG 3,350
5	0.2 M Magnesium chloride hexahydrate	20% w/v PEG 3,350
6	0.2 M Sodium chloride	20% w/v PEG 3,350
7	0.2 M Calcium chloride dehydrate	20% w/v PEG 3,350
8	0.2 M Potassium chloride	20% w/v PEG 3,350
9	0.2 M Ammonium chloride	20% w/v PEG 3,350
10	0.2 M Sodium iodide	20% w/v PEG 3,350
11	0.2 M Potassium iodide	20% w/v PEG 3,350
12	0.2 M Ammonium iodide	20% w/v PEG 3,350
13	0.2 M Sodium thiocyanate	20% w/v PEG 3,350
14	0.2 M Potassium thiocyanate	20% w/v PEG 3,350
15	0.2 M Lithium nitrate	20% w/v PEG 3,350
16	0.2 M Magnesium nitrate hexahydrate	20% w/v PEG 3,350
17	0.2 M Sodium nitrate	20% w/v PEG 3,350

18	0.2 M Potassium nitrate	20% w/v PEG 3,350
19	0.2 M Ammonium nitrate	20% w/v PEG 3,350
20	0.2 M Magnesium formate dihydrate	20% w/v PEG 3,350
21	0.2 M Sodium formate	20% w/v PEG 3,350
22	0.2 M Potassium formate	20% w/v PEG 3,350
23	0.2 M Ammonium formate	20% w/v PEG 3,350
24	0.2 M Lithium acetate dihydrate	20% w/v PEG 3,350
25	0.2 M Magnesium acetate tetrahydrate	20% w/v PEG 3,350
26	0.2 M Zinc acetate dihydrate	20% w/v PEG 3,350
27	0.2 M Sodium acetate trihydrate	20% w/v PEG 3,350
28	0.2 M Calcium acetate hydrate	20% w/v PEG 3,350
29	0.2 M Potassium acetate	20% w/v PEG 3,350
30	0.2 M Ammonium acetate	20% w/v PEG 3,350
31	0.2 M Lithium sulfate monohydrate	20% w/v PEG 3,350
32	0.2 M Magnesium sulfate heptahydrate	20% w/v PEG 3,350
33	0.2 M Sodium sulfate decahydrate	20% w/v PEG 3,350
34	0.2 M Potassium sulfate	20% w/v PEG 3,350
35	0.2 M Ammonium sulfate	20% w/v PEG 3,350
36	0.2 M Sodium tartrate dibasic dihydrate	20% w/v PEG 3,350
37	0.2 M Potassium sodium tartrate tetrahydrate	20% w/v PEG 3,350
38	0.2 M Ammonium tartrate dibasic	20% w/v PEG 3,350
39	0.2 M Sodium phosphate	20% w/v PEG 3,350

	monobasic monohydrate	
40	0.2 M Sodium phosphate dibasic dihydrate	20% w/v PEG 3,350
41	0.2 M Potassium phosphate monobasic	20% w/v PEG 3,350
42	0.2 M Potassium phosphate dibasic	20% w/v PEG 3,350
43	0.2 M Ammonium phosphate monobasic	20% w/v PEG 3,350
44	0.2 M Ammonium phosphate dibasic	20% w/v PEG 3,350
45	0.2 M Lithium citrate tribasic tetrahydrate	20% w/v PEG 3,350
46	0.2 M Sodium citrate tribasic dihydrate	20% w/v PEG 3,350
47	0.2 M Potassium citrate tribasic monohydrate	20% w/v PEG 3,350
48	0.2 M Ammonium citrate dibasic	20% w/v PEG 3,350

Table J2 Hampton Crystal Screen I components

Tube No	Salt	Buffer	Precipitant
1	0.02 M Calcium chloride dihydrate	0.1 M Sodium acetate trihydrate pH 4.6	30% (v/v) 2-Methyl-2,4-pentanediol
2	None	None	0.4 M Potassium sodium tartrate tetrahydrate
3	None	None	0.4 M Ammonium phosphate monobasic
4	None	0.1 M TRIS hydrochloride pH	2.0 M Ammonium sulfate

		8.5	
5	0.2 M Sodium citrate tribasic dihydrate	0.1 M HEPES sodium pH 7.5	30% (v/v) 2-Methyl-2,4-pentanediol
6	0.2 M Magnesium chloride hexahydrate	0.1 M TRIS hydrochloride pH 8.5	30% (w/v) Polyethylene glycol 4,000
7	None	0.1 M Sodium cacodylate trihydrate pH 6.5	1.4 M Sodium acetate trihydrate
8	0.2 M Sodium citrate tribasic dihydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% (v/v) 2-Propanol
9	0.2 M Ammonium acetate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	30% (w/v) Polyethylene glycol 4,000
10	0.2 M Ammonium acetate	0.1 M Sodium acetate trihydrate pH 4.6	30% (w/v) Polyethylene glycol 4,000
11	None	0.1 M Sodium citrate tribasic dihydrate pH 5.6	1.0 M Ammonium phosphate monobasic
12	0.2 M Magnesium chloride hexahydrate	0.1 M HEPES sodium pH 7.5	30% (v/v) 2-Propanol
13	0.2 M Sodium citrate tribasic dihydrate	0.1 M TRIS hydrochloride pH 8.5	30% (v/v) Polyethylene glycol 400
14	0.2 M Calcium	0.1 M HEPES sodium pH 7.5	28% (v/v) Polyethylene glycol 400

	chloride dihydrate		
15	0.2 M Ammonium sulfate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% (w/v) Polyethylene glycol 8,000
16	None	0.1 M HEPES sodium pH 7.5	1.5 M Lithium sulfate monohydrate
17	0.2 M Lithium sulfate monohydrate	0.1 M TRIS hydrochloride pH 8.5	30% (w/v) Polyethylene glycol 4,000
18	0.2 M Magnesium acetate tetrahydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	20% (w/v) Polyethylene glycol 8,000
19	0.2 M Ammonium acetate	0.1 M TRIS hydrochloride pH 8.5	30% (v/v) 2-Propanol
20	0.2 M Ammonium sulfate	0.1 M Sodium acetate trihydrate pH 4.6	25% (w/v) Polyethylene glycol 4,000
21	0.2 M Magnesium acetate tetrahydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% (v/v) 2-Methyl-2,4-pentanediol
22	0.2 M Sodium acetate trihydrate	0.1 M TRIS hydrochloride pH 8.5	30% (w/v) Polyethylene glycol 4,000
23	0.2 M Magnesium chloride hexahydrate	0.1 M HEPES sodium pH 7.5	30% (v/v) Polyethylene glycol 400
24	0.2 M Calcium chloride dihydrate	0.1 M Sodium acetate trihydrate pH 4.6	20% (v/v) 2-Propanol

25	None	0.1 M Imidazole pH 6.5	1.0 M Sodium acetate trihydrate
26	0.2 M Ammonium acetate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	30% (v/v) 2-Methyl-2,4-pentanediol
27	0.2 M Sodium citrate tribasic dihydrate	0.1 M HEPES sodium pH 7.5	20% (v/v) 2-Propanol
28	0.2 M Sodium acetate trihydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	30% (w/v) Polyethylene glycol 8,000
29	None	0.1 M HEPES sodium pH 7.5	0.8 M Potassium sodium tartrate tetrahydrate
30	0.2 M Ammonium sulfate	None	30% (w/v) Polyethylene glycol 8,000
31	0.2 M Ammonium sulfate	None	30% (w/v) Polyethylene glycol 4,000
32	None	None	2.0 M Ammonium sulphate
33	None	None	4.0 M Sodium formate
34	None	0.1 M Sodium acetate trihydrate pH 4.6	2.0 M Sodium formate
35	None	0.1 M HEPES sodium pH 7.5	0.8 M Sodium phosphate monobasic monohydrate, 0.8 MPotassium phosphate monobasic
36	None	0.1 M TRIS hydrochloride pH 8.5	8% (w/v) Polyethylene glycol 8,000
37	None	0.1 M Sodium acetate trihydrate	8% (w/v) Polyethylene glycol

		pH 4.6	8,000
38	None	0.1 M HEPES sodium pH 7.5	1.4 M Sodium citrate tribasic dihydrate
39	None	0.1 M HEPES sodium pH 7.5	2% (v/v) Polyethylene glycol 400, 2.0 M Ammonium sulfate
40	None	0.1 M Sodium citrate tribasic dihydrate pH 5.6	20% (v/v) 2-Propanol, 20% (w/v) Polyethylene glycol 4,000
41	None	0.1 M HEPES sodium pH 7.5	10% (v/v) 2-Propanol, 20% (w/v) Polyethylene glycol 4,000
42	0.05 M Potassium phosphate monobasic	None	20% (w/v) Polyethylene glycol 8,000
43	None	None	30% (w/v) Polyethylene glycol 1,500
44	None	None	0.2 M Magnesium formate dihydrate
45	0.2 M Zinc acetate dihydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	18% (w/v) Polyethylene glycol 8,000
46	0.2 M Calcium acetate hydrate	0.1 M Sodium cacodylate trihydrate pH 6.5	18% (w/v) Polyethylene glycol 8,000
47	None	0.1 M Sodium acetate trihydrate pH 4.6	2.0 M Ammonium sulfate
48	None	0.1 M TRIS hydrochloride pH 8.5	2.0 M Ammonium phosphate monobasic

Table J3 Hampton Crystal Screen II components

Tube No	Salt	Buffer	Precipitant
1	2.0 M Sodium chloride	None	10% (w/v) Polyethylene glycol 6,000
2	0.5 M Sodium chloride, 0.01 M Magnesium chloride hexahydrate	None	0.01M Hexadecyltrimethylammonium bromide
3	None	None	25% (v/v) Ethylene glycol
4	None	None	35% (v/v) 1,4-Dioxane
5	2.0 M Ammonium sulfate	None	5% (v/v) 2-Propanol
6	None	None	1.0 M Imidazole pH 7.0
7	None	None	10% (w/v) Polyethylene glycol 1,000, 10% (w/v) Polyethylene glycol 8,000
8	1.5 M Sodium chloride	None	10% (v/v) Ethanol
9	None	0.1 M Sodium acetate trihydrate pH 4.6	2.0 M Sodium chloride
10	0.2 M Sodium chloride	0.1 M Sodium acetate trihydrate pH 4.6	30% (v/v) 2-Methyl-2,4-pentanediol
11	0.01 M Cobalt(II) chloride hexahydrate	0.1 M Sodium acetate trihydrate pH 4.6	1.0 M 1,6-Hexanediol
12	0.1 M Cadmium chloride hydrate	0.1 M Sodium acetate trihydrate pH 4.6	30% (v/v) Polyethylene glycol 400

13	0.2 M Ammonium sulfate	0.1 M Sodium acetate trihydrate pH 4.6	30% (w/v) Polyethylene glycol monomethyl ether 2,000
14	0.2 M Potassium sodium tartrate tetrahydrate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	2.0 M Ammonium sulfate
15	0.5 M Ammonium sulfate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	1.0 M Lithium sulfate monohydrate
16	0.5 M Sodium chloride	0.1 M Sodium citrate tribasic dihydrate pH 5.6	2% (v/v) Ethylene imine polymer
17	None	0.1 M Sodium citrate tribasic dihydrate pH 5.6	35% (v/v) tert-Butanol
18	0.01 M Iron(III) chloride hexahydrate	0.1 M Sodium citrate tribasic dihydrate pH 5.6	10% (v/v) Jeffamine M-600
19	None	0.1 M Sodium citrate tribasic dihydrate pH 5.6	2.5 M 1,6-Hexanediol
20	None	0.1 M MES monohydrate pH 6.5	1.6 M Magnesium sulfate heptahydrate
21	0.1 M Sodium phosphate monobasic monohydrate, 0.1 M Potassium phosphate monobasic	0.1 M MES monohydrate pH 6.5	2.0 M Sodium chloride
22	None	0.1 M MES monohydrate pH 6.5	12% (w/v) Polyethylene glycol 20,000
23	1.6 M Ammonium	0.1 M MES monohydrate	10% (v/v) 1,4-Dioxane

	sulfate	pH 6.5	
24	0.05 M Cesium chloride	0.1 M MES monohydrate pH 6.5	30% (v/v) Jeffamine M-600
25	0.01 M Cobalt(II) chloride hexahydrate	0.1 M MES monohydrate pH 6.5	1.8 M Ammonium sulfate
26	0.2 M Ammonium sulfate	0.1 M MES monohydrate pH 6.5	30% (w/v) Polyethylene glycol monomethyl ether 5,000
27	0.01 M Zinc sulfate heptahydrate	0.1 M MES monohydrate pH 6.5	25% (v/v) Polyethylene glycol monomethyl ether 550
28	None	None	1.6 M Sodium citrate tribasic dihydrate pH 6.5
29	0.5 M Ammonium sulfate	0.1 M HEPES pH 7.5	30% (v/v) 2-Methyl-2,4-pentanediol
30	None	0.1 M HEPES pH 7.5	10% (w/v) Polyethylene glycol 6,000, 5% (v/v) 2-Methyl-2,4-pentanediol
31	None	0.1 M HEPES pH 7.5	20% (v/v) Jeffamine M-600
32	0.1 M Sodium chloride	0.1 M HEPES pH 7.5	1.6 M Ammonium sulfate
33	None	0.1 M HEPES pH 7.5	2.0 M Ammonium formate
34	0.05 M Cadmium sulfate hydrate	0.1 M HEPES pH 7.5	1.0 M Sodium acetate trihydrate
35	None	0.1 M HEPES pH 7.5	70% (v/v) 2-Methyl-2,4-pentanediol
36	None	0.1 M HEPES pH 7.5	4.3 M Sodium chloride
37	None	0.1 M HEPES pH 7.5	10% (w/v) Polyethylene glycol 8,000, 8% (v/v) Ethylene glycol
38	None	0.1 M HEPES pH 7.5	20% (w/v) Polyethylene glycol 10,000
39	0.2 M Magnesium chloride	0.1 M Tris pH 8.5	3.4 M 1,6-Hexanediol

	hexahydrate		
40	None	0.1 M Tris pH 8.5	25% (v/v) tert-Butanol
41	0.01 M Nickel(II) chloride hexahydrate	0.1 M Tris pH 8.5	1.0 M Lithium sulfate monohydrate
42	1.5 M Ammonium sulfate	0.1 M Tris pH 8.5	12% (v/v) Glycerol
43	0.2 M Ammonium phosphate monobasic	0.1 M Tris pH 8.5	50% (v/v) 2-Methyl-2,4-pentanediol
44	None	0.1 M Tris pH 8.5	20% (v/v) Ethanol
45	0.01 M Nickel(II) chloride hexahydrate	0.1 M Tris pH 8.5	20% (w/v) Polyethylene glycol monomethyl ether 2,000
46	0.1 M Sodium chloride	0.1 M BICINE pH 9.0	20% (v/v) Polyethylene glycol monomethyl ether 550
47	None	0.1 M BICINE pH 9.0	2.0 M Magnesium chloride hexahydrate
48	None	0.1 M BICINE pH 9.0	2% (v/v) 1,4-Dioxane, 10% (w/v) Polyethylene glycol 20,000

Table J4 Hampton SaltRx Screen components

Tube No	Salt	Buffer
1	1.8 M Sodium acetate pH 7.0	0.1 M Bis-Tris propane pH 7.0
2	2.8 M Sodium acetate pH 7.0	0.1 M Bis-Tris propane pH 7.0
3	1.5 M Ammonium chloride	0.1 M Sodium acetate pH 4.6
4	1.5 M Ammonium chloride	0.1 M Bis-Tris propane pH 7.0
5	1.5 M Ammonium chloride	0.1 M Tris pH 8.5
6	3.5 M Ammonium chloride	0.1 M Sodium acetate pH 4.6
7	3.5 M Ammonium chloride	0.1 M Bis-Tris propane pH 7.0
8	3.5 M Ammonium chloride	0.1 M Tris pH 8.5
9	2.2 M Sodium chloride	0.1 M Sodium acetate pH 4.6
10	2.2 M Sodium chloride	0.1 M BIS-TRIS propane pH 7.0
11	2.2 M Sodium chloride	0.1 M Tris pH 8.5
12	3.2 M Sodium chloride	0.1 M Sodium acetate trihydrate pH 4.6
13	3.2 M Sodium chloride	M BIS-TRIS propane pH 7.0
14	3.2 M Sodium chloride	0.1 M Tris pH 8.5
15	1.0 M Ammonium citrate dibasic	0.1 M Sodium acetate trihydrate pH 4.6

16	1.8 M Ammonium citrate dibasic	0.1 M Sodium acetate trihydrate pH 4.6
17	1.0 M Ammonium citrate tribasic pH 7.0	0.1 M BIS-TRIS propane pH 7.0
18	2.0 M Ammonium citrate tribasic pH 7.0	0.1 M BIS-TRIS propane pH 7.0
19	0.7 M Sodium citrate tribasic dihydrate	0.1 M BIS-TRIS propane pH 7.0
20	0.7 M Sodium citrate tribasic dihydrate	0.1 M Tris pH 8.5
21	1.2 M Sodium citrate tribasic dihydrate	0.1 M BIS-TRIS propane pH 7.0
22	1.2 M Sodium citrate tribasic dihydrate	0.1 M Tris pH 8.5
23	0.4 M Magnesium formate dihydrate	0.1 M Sodium acetate trihydrate pH 4.6
24	0.4 M Magnesium formate dihydrate	0.1 M BIS-TRIS propane pH 7.0
25	0.4 M Magnesium formate dihydrate	0.1 M Tris pH 8.5
26	0.7 M Magnesium formate dihydrate	0.1 M BIS-TRIS propane pH 7.0
27	2.0 M Sodium formate	0.1 M Sodium acetate trihydrate pH 4.6
28	2.0 M Sodium formate	0.1 M BIS-TRIS propane pH 7.0
29	2.0 M Sodium formate	0.1 M Tris pH 8.5
30	3.5 M Sodium formate	0.1 M Sodium acetate trihydrate pH 4.6
31	3.5 M Sodium formate	0.1 M BIS-TRIS propane pH 7.0
32	3.5 M Sodium formate	0.1 M Tris pH 8.5
33	1.2 M DL-Malic acid pH 7.0	0.1 M BIS-TRIS propane pH 7.0
34	2.2 M DL-Malic acid	0.1 M BIS-TRIS propane pH 7.0

	pH 7.0	
35	1.4 M Sodium malonate pH 7.0	0.1 M BIS-TRIS propane pH 7.0
36	2.4 M Sodium malonate pH 7.0	0.1 M BIS-TRIS propane pH 7.0
37	2.5 M Ammonium nitrate	0.1 M Sodium acetate trihydrate pH 4.6
38	2.5 M Ammonium nitrate	0.1 M BIS-TRIS propane pH 7.0
39	2.5 M Ammonium nitrate	0.1 M Tris pH 8.5
40	6.0 M Ammonium nitrate	0.1 M Sodium acetate trihydrate pH 4.6
41	6.0 M Ammonium nitrate	0.1 M BIS-TRIS propane pH 7.0
42	6.0 M Ammonium nitrate	0.1 M Tris pH 8.5
43	1.5 M Sodium nitrate	0.1 M Sodium acetate trihydrate pH 4.6
44	1.5 M Sodium nitrate	0.1 M BIS-TRIS propane pH 7.0
45	1.5 M Sodium nitrate	0.1 M Tris pH 8.5
46	4.0 M Sodium nitrate	0.1 M Sodium acetate trihydrate pH 4.6
47	4.0 M Sodium nitrate	0.1 M BIS-TRIS propane pH 7.0
48	4.0 M Sodium nitrate	0.1 M Tris pH 8.5
49	1.0 M Ammonium phosphate monobasic	0.1 M Sodium acetate trihydrate pH 4.6
50	1.8 M Ammonium phosphate monobasic	0.1 M Sodium acetate trihydrate pH 4.6
51	1.5 M Ammonium phosphate dibasic	0.1 M Tris pH 8.5
52	2.4 M Ammonium phosphate dibasic	0.1 M Tris pH 8.5

53	1.0 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 5.0	None
54	1.0 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 6.9	None
55	1.0 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 8.2	None
56	1.8 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 5.0	None
57	1.8 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 6.9	None
58	1.8 M Sodium phosphate monobasic monohydrate, Potassium phosphate dibasic / pH 8.2	None
59	0.5 M Succinic acid pH 7.0	0.1 M BIS-TRIS propane pH 7.0
60	1.0 M Succinic acid pH 7.0	0.1 M BIS-TRIS propane pH 7.0
61	1.5 M Ammonium sulfate	0.1 M Sodium acetate trihydrate pH 4.6
62	1.5 M Ammonium sulfate	0.1 M BIS-TRIS propane pH 7.0
63	1.5 M Ammonium sulfate	0.1 M Tris pH 8.5

64	2.5 M Ammonium sulfate	0.1 M Sodium acetate trihydrate pH 4.6
65	2.5 M Ammonium sulfate	0.1 M BIS-TRIS propane pH 7.0
66	2.5 M Ammonium sulfate	0.1 M Tris pH 8.5
67	0.8 M Lithium sulfate monohydrate	0.1 M Sodium acetate trihydrate pH 4.6
68	0.8 M Lithium sulfate monohydrate	0.1 M BIS-TRIS propane pH 7.0
69	0.8 M Lithium sulfate monohydrate	0.1 M Tris pH 8.5
70	1.5 M Lithium sulfate monohydrate	0.1 M Sodium acetate trihydrate pH 4.6
71	1.5 M Lithium sulfate monohydrate	0.1 M BIS-TRIS propane pH 7.0
72	1.5 M Lithium sulfate monohydrate	0.1 M Tris pH 8.5
73	1.0 M Magnesium sulfate hydrate	0.1 M Sodium acetate trihydrate pH 4.6
74	1.0 M Magnesium sulfate hydrate	0.1 M BIS-TRIS propane pH 7.0
75	1.0 M Magnesium sulfate hydrate	0.1 M Tris pH 8.5
76	1.8 M Magnesium sulfate hydrate	0.1 M Sodium acetate trihydrate pH 4.6
77	1.8 M Magnesium sulfate hydrate	0.1 M BIS-TRIS propane pH 7.0
78	1.8 M Magnesium sulfate hydrate	0.1 M Tris pH 8.5
79	0.7 M Ammonium tartrate dibasic	0.1 M Sodium acetate trihydrate pH 4.6
80	0.7 M Ammonium tartrate dibasic	0.1 M BIS-TRIS propane pH 7.0
81	0.7 M Ammonium tartrate dibasic	0.1 M Tris pH 8.5
82	1.0 M Ammonium tartrate dibasic	0.1 M Sodium acetate trihydrate pH 4.6
83	1.3 M Ammonium tartrate dibasic	0.1 M BIS-TRIS propane pH 7.0
84	1.4 M Ammonium tartrate dibasic	0.1 M Tris pH 8.5

85	0.6 M Potassium sodium tartrate tetrahydrate	0.1 M BIS-TRIS propane pH 7.0
86	1.2 M Potassium sodium tartrate tetrahydrate	0.1 M BIS-TRIS propane pH 7.0
87	0.6 M Potassium sodium tartrate tetrahydrate	0.1 M Tris pH 8.5
88	1.2 M Potassium sodium tartrate tetrahydrate	0.1 M Tris pH 8.5
89	0.5 M Potassium thiocyanate	0.1 M Sodium acetate trihydrate pH 4.6
90	0.5 M Potassium thiocyanate	0.1 M BIS-TRIS propane pH 7.0
91	0.5 M Potassium thiocyanate	0.1 M Tris pH 8.5
92	4.0 M Ammonium acetate	0.1 M Sodium acetate trihydrate pH 4.6
93	4.0 M Ammonium acetate	0.1 M BIS-TRIS propane pH 7.0
94	4.0 M Ammonium acetate	0.1 M Tris pH 8.5
95	35% v/v Tacsimate pH 7.0	0.1 M BIS-TRIS propane pH 7.0
96	60% v/v Tacsimate pH 7.0	0.1 M BIS-TRIS propane pH 7.0

Table J5: Molecular Dimensions Clear Strategy Screen I

Tube No	**Salt**	**Precipitant**
1	0.3 M Sodium acetate trihydrate	25 % PEG 2000 monomethyl ether
2	0.2 M Lithium sulphate	25 % PEG 2000 monomethyl ether
3	0.2 M Magnesium chloride	25 % PEG 2000 monomethyl ether

4	0.2 M Potassium bromide	25 % PEG 2000 monomethyl ether
5	0.2 M Potassium thiocyanate	25 % PEG 2000 monomethyl ether
6	0.8 M Sodium formate	25 % PEG 2000 monomethyl ether
7	0.3 M Sodium acetate	15 % PEG 4000
8	0.2 M Lithium sulphate	15 % PEG 4000
9	0.2 M Magnesium chloride	15 % PEG 4000
10	0.2 M Potassium bromide	15 % PEG 4000
11	0.2 M Potassium thiocyanate	15 % PEG 4000
12	0.8 M Sodium formate	15 % PEG 4000
13	0.3 M Sodium acetate trihydrate	10 % PEG 1000
14	0.2 M Lithium sulphate	10 % PEG 1000
15	0.2 M Magnesium chloride	10 % PEG 1000
16	0.2 M Potassium bromide	10 % PEG 1000
17	0.2 M Potassium thiocyanate	10 % PEG 1000
18	0.8 M Sodium formate	10 % PEG 1000
19	0.3 M Sodium acetate trihydrate	8 % PEG 20000 + 8 % PEG 500 monomethyl ether

20	0.2 M Lithium sulphate	8 % PEG 20000 + 8 % PEG 500 monomethyl ether
21	0.2 M Magnesium chloride	8 % PEG 20000 + 8 % PEG 500 monomethyl ether
22	0.2 M Potassium bromide	8 % PEG 20000 + 8 % PEG 500 monomethyl ether
23	0.2 M Potassium thiocyanate	8 % PEG 20000 + 8 % PEG 500 monomethyl ether
24	0.8 M Sodium formate	8 % PEG 20000 + 8 % PEG 500 monomethyl ether

Table J6: Molecular Dimensions Clear Strategy Screen II

Tube No	**Salt**	**Precipitant**
1	1.5 M Ammonium sulphate	None
2	0.8 M Lithium sulphate	None
3	2 M Sodium formate	None
4	0.5 M Potassium phosphate monobasic	None
5	0.2 M Calcium acetate	25 % PEG 2000 monomethyl ether
6	0.2 M Calcium acetate	15 % PEG 4000
7	2.7 M Ammonium sulphate	None
8	1.8 M Lithium sulphate	None
9	4 M Sodium formate	None

10	1.0 M Potassium phosphate monobasic	None
11	0.2 M Calcium acetate	10 % PEG 8000 + 10 % PEG 1000
12	0.2 M Calcium acetate	8 % PEG 20000 + 8 % PEG 500 monomethyl ether
13	None	40 % 2-methyl-2,4-pentanediol
14	None	40 % 1,4-butanediol
15	5 mM Cadmium chloride	20 % PEG 4000
16	0.15 M Potassium thiocyanate	20 % PEG 500 monomethyl ether
17	0.15 M Potassium thiocyanate	20 % PEG 600
18	0.15 M Potassium thiocyanate	20% PEG 1500
19	None	35 % v/v propanol
20	None	30 % v/v Jeffamine
21	5 mM Nickel chloride	20 % PEG 4000
22	0.15 M Potassium thiocyanate	18 % PEG 3350
23	0.15 M Potassium thiocyanate	18 % PEG 5000 monomethyl ether
24	0.15 M Potassium thiocyanate	15 % PEG 6000

Appendix K

Sequencing Results

GENE ID: 1238265 vapA | virulence associated protein VapA [*Rhodococcus equi*] (10 or fewer PubMed links)

```
 Score =   283 bits (724),  Expect = 1e-98, Method: Compositional matrix
adjust.
 Identities = 158/161 (98%), Positives = 160/161 (99%), Gaps = 0/161 (0%)
 Frame = +2

Query  104  SHMTVLDSGSSSAILNSGAGSGIVGSGSYDSSTTSLNLQKDEPNGRASDTAGQEQQYDVH
283
            ++ TVLDSGSSSAILNSGAGSGIVGSGSYDSSTTSLNLQKDEPNGRASDTAGQEQQYDVH
Sbjct  29   ANATVLDSGSSSAILNSGAGSGIVGSGSYDSSTTSLNLQKDEPNGRASDTAGQEQQYDVH  88

Query  284  GDVISAVVYQRFHVFGPEGKVFDGDAGGLTLPGAGAFWGTLFTNDLQRLYKDTVSFQYNA
463
            GDVISAVVYQRFHVFGPEGKVFDGDAGGLTLPGAGAFWGTLFTNDLQRLYKDTVSFQYNA
Sbjct  89   GDVISAVVYQRFHVFGPEGKVFDGDAGGLTLPGAGAFWGTLFTNDLQRLYKDTVSFQYNA
148

Query  464  VGPYLNINFFDSSGSFLGHIQSGGVSTVVGVGGGSGSWHNA  586
            VGPYLNINFFDSSGSFLGHIQSGGVSTVVGVGGGSGSWHNA
Sbjct  149  VGPYLNINFFDSSGSFLGHIQSGGVSTVVGVGGGSGSWHNA  189
```

GENE ID: 1238267 vapC | virulence associated protein VapC [*Rhodococcus equi*] (10 or fewer PubMed links)

```
 Score =   295 bits (755),  Expect = 8e-104, Method: Compositional matrix
adjust.
 Identities = 145/148 (98%), Positives = 146/148 (99%), Gaps = 0/148 (0%)
 Frame = +2

Query  110  SHMNVVAPSAWGGAQSAADKEGEGVTLGGVGVLRPHNKDADEQYTVHGVVVSALFYNHLR
289
            +  NVVAPSAWGGAQSAADKEGEGVTLGGVGVLRPHNKDADEQYTVHGVVVSALFYNHLR
Sbjct  27   ARANVVAPSAWGGAQSAADKEGEGVTLGGVGVLRPHNKDADEQYTVHGVVVSALFYNHLR  86

Query  290  ISVDGGMTFDGDGGGLSTPGGGALWGTLTTSDLQQLYDETASFECNAVGPYLNINFYDSY
469
            ISVDGGMTFDGDGGGLSTPGGGALWGTLTTSDLQQLYDETASFECNAVGPYLNINFYDSY
Sbjct  87   ISVDGGMTFDGDGGGLSTPGGGALWGTLTTSDLQQLYDETASFECNAVGPYLNINFYDSY
146

Query  470  GRILASVQAGGVSTMIGIGGGNGRWHLV  553
            GRILASVQAGGVSTMIGIGGGNGRWHLV
Sbjct  147  GRILASVQAGGVSTMIGIGGGNGRWHLV  174
```

GENE ID: 1238269 vapE | virulence associated protein VapE [*Rhodococcus equi*] (10 or fewer PubMed links)

Score = 328 bits (841), Expect = 3e-113, Method: Compositional matrix adjust.
Identities = 160/160 (100%), Positives = 160/160 (100%), Gaps = 0/160 (0%)
Frame = +1

```
Query  118  REIGAQAWPASQLESGLAVSGNPVGVHDVRMAVHDDSTHTREFKEDDSEKQYPVHGFASS
297
            REIGAQAWPASQLESGLAVSGNPVGVHDVRMAVHDDSTHTREFKEDDSEKQYPVHGFASS
Sbjct  47   REIGAQAWPASQLESGLAVSGNPVGVHDVRMAVHDDSTHTREFKEDDSEKQYPVHGFASS
106

Query  298  FIFYQTVSIIIDDDGRGGPGKTFEGEAGGITTPGAAGYAGVLFTSDLERLYRETVSFEYN
477
            FIFYQTVSIIIDDDGRGGPGKTFEGEAGGITTPGAAGYAGVLFTSDLERLYRETVSFEYN
Sbjct  107  FIFYQTVSIIIDDDGRGGPGKTFEGEAGGITTPGAAGYAGVLFTSDLERLYRETVSFEYN
166

Query  478  AVGPYLNINLFAGDGGLLGHVQSGAISSLVGIGGGTGAWR  597
            AVGPYLNINLFAGDGGLLGHVQSGAISSLVGIGGGTGAWR
Sbjct  167  AVGPYLNINLFAGDGGLLGHVQSGAISSLVGIGGGTGAWR  206
```

GENE ID: 1238341 VapG |virulence associated protein VapG [*Rhodococcus equi*]

```
Query  88   SGLVPRG-SHMETSMVSTTAASSVEHAANTYDFAEAKSGSSIPAKVAAEQANSYSVHGLV
264
            S L P G ++ ETSMVSTTAASSVEHAANTYDFAEAKSGSSIPAKVAAEQANSYSVHGLV
Sbjct  16   SVLAPAGIANAETSMVSTTAASSVEHAANTYDFAEAKSGSSIPAKVAAEQANSYSVHGLV  75

Query  265  TSLAVYQHFSLTVEGGGKTFTGDSGGISIPGVAVLEGTLFTEDLQHLYSDTVSFEYNAVG
444
            TSLAVYQHFSLTVEGGGKTFTGDSGGISIPGVAVLEGTLFTEDLQHLYSDTVSFEYNAVG
Sbjct  76   TSLAVYQHFSLTVEGGGKTFTGDSGGISIPGVAVLEGTLFTEDLQHLYSDTVSFEYNAVG
135

Query  445  PYLNINFFDSHGTLLGHVQSGSIGTVSGIGGGTGGWQ  555
            PYLNINFFDSHGTLLGHVQSGSIGTVSGIGGGTGGWQ
Sbjct  136  PYLNINFFDSHGTLLGHVQSGSIGTVSGIGGGTGGWQ  172
```

GENE ID: 1238264 vapH | virulence associated protein VapH [*Rhodococcus equi*] (10 or fewer PubMed links)

Score = 320 bits (820), Expect = 2e-113, Method: Compositional matrix adjust.
Identities = 159/163 (98%), Positives = 159/163 (98%), Gaps = 1/163 (1%)
Frame = +2

```
Query  95   VPRGSHMAPPPLPDGPTHDLPTWREEGANYSDGTMLVRASSNFLEPSTHSDSGQQQWTVQ
274
            VP G   APPPLPDGPTHDLPTWREEGANYSDGTMLVRASSNFLEPSTHSDSGQQQWTVQ
Sbjct  26   VPWGC-AAPPPLPDGPTHDLPTWREEGANYSDGTMLVRASSNFLEPSTHSDSGQQQWTVQ  84

Query  275  GVLASALVYQRLKLNVEGGETFEGYAGGLSFPGGAMVWGTLFTDNIQRLYDRTESFEFNA
454
```

```
            GVLASALVYQRLKLNVEGGETFEGYAGGLSFPGGAMVWGTLFTDNIQRLYDRTESFEFNA
Sbjct  85   GVLASALVYQRLKLNVEGGETFEGYAGGLSFPGGAMVWGTLFTDNIQRLYDRTESFEFNA
144

Query  455  VGPYLNVNFFDGHSAILGHAQLGGVSSVIGIGGGTGTWIGDVA  583
            VGPYLNVNFFDGHSAILGHAQLGGVSSVIGIGGGTGTWIGDVA
Sbjct  145  VGPYLNVNFFDGHSAILGHAQLGGVSSVIGIGGGTGTWIGDVA  187
```

pET23-cmt

esterase [*Corynebacterium glutamicum* ATCC 13032]
Sequence ID: gi|19551592|ref|NP_599594.1|Length: 365Number of Matches: 1
Gene-associated gene details
Identical Proteins-Proteins identical to the subject
Range 1: 30 to 365GenPeptGraphics Next Match Previous Match First Match

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
660 bits(1702)	0.0()	Compositional matrix adjust.	324/336(96%)	324/336(96%)	0/336(0%)	+2

```
Query  56    AEVTPADVAGDTALSTISDSAPADEASAPRWRAHVNAADERVKEMWAYSPSMDRNVPLVV
235

             AEVTPADVAGDTALSTISDSAPADEASAPRWRAHVNAADERVKEMWAYSPSMDRNVPLVV
Sbjct  30    AEVTPADVAGDTALSTISDSAPADEASAPRWRAHVNAADERVKEMWAYSPSMDRNVPLVV
89

Query  236   ITADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPMEGKFSYYTDWV
415

             ITADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPMEGKFSYYTDWV
Sbjct  90    ITADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPMEGKFSYYTDWV
149

Query  416   EENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLFPQHFPGFYDAA
595

             EENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLFPQHFPGFYDAA
Sbjct  150   EENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLFPQHFPGFYDAA
209

Query  596   ASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALINSDKLRGTELYV
775

             ASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALINSDKLRGTELYV
Sbjct  210   ASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALINSDKLRGTELYV
269

Query  776   SNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTHDLKAKLDSAGI
955

             SNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTHDLKAKLDSAGI
Sbjct  270   SNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTHDLKAKLDSAGI
329

Query  956   PADWKPPPNRHPLMGLVADDLRGSWTTFARAFELEA  1063
             PADW    P      G   DDLRGSWTTFARAFELEA
Sbjct  330   PADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA  365
```

Rev

Range 1: 30 to 365GenPeptGraphics Next Match Previous Match First Match

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
588 bits(1515)	0.0()	Compositional matrix adjust.	296/347(85%)	301/347(86%)	21/347(6%)	+3

Features:

```
Query  210   AEVTPADVAGDTALIHHLR*CSCR*SLCTRWR-RTRQRSRRARQ---------RNVGILP
359

             AEVTPADVAGDTAL           +   RWR        R ++          RNV ++
Sbjct  30    AEVTPADVAGDTALSTISDSAPADEASAPRWRAHVNAADERVKEMWAYSPSMDRNVPLVV
89

Query  360   FHGTAMCHW*L*TADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPM
539

                         TADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPM
Sbjct  90    I-----------TADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPM
138

Query  540   EGKFSYYTDWVEENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLF
719

             EGKFSYYTDWVEENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLF
Sbjct  139   EGKFSYYTDWVEENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLF
198

Query  720   PQHFPGFYDAAASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALIN
899

             PQHFPGFYDAAASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALIN
Sbjct  199   PQHFPGFYDAAASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALIN
258

Query  900   SDKLRGTELYVSNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTH
1079

             SDKLRGTELYVSNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTH
Sbjct  259   SDKLRGTELYVSNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTH
318

Query  1080  DLKAKLDSAGIPADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA  1220
             DLKAKLDSAGIPADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA
Sbjct  319   DLKAKLDSAGIPADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA  365
```

pET28-cmt

For

Related Information
Identical Proteins-Proteins identical to the subject
Range 1: 1 to 243GenPeptGraphics Next Match Previous Match First Match

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
443 bits(1139)	4e-170(2)	Compositional matrix adjust.	219/243(90%)	224/243(92%)	0/243(0%)	+2

Features:

```
Query  35    LTLRRRYTMGSSHHHHHHSSGLVPRGSHMAEVTPADVAGDTALSTISDSAPADEASAPRW
214

             + L RR    +       S+ + P  +  AEVTPADVAGDTALSTISDSAPADEASAPRW
Sbjct  1     MKLLRRIAAPAIALGIAMSTIVTPSTAGAAEVTPADVAGDTALSTISDSAPADEASAPRW  60
```

```
Query  215  RAHVNAADERVKEMWAYSPSMDRNVPLVVITADESAGPRPVIYLLNGGDGGEGAANWVMQ
394

             RAHVNAADERVKEMWAYSPSMDRNVPLVVITADESAGPRPVIYLLNGGDGGEGAANWVMQ
Sbjct  61    RAHVNAADERVKEMWAYSPSMDRNVPLVVITADESAGPRPVIYLLNGGDGGEGAANWVMQ
120

Query  395  TDVLDFYLEKNVNVVIPMEGKFSYYTDWVEENASLGGKQMWETFLVKELPGPLEEKLNTD
574

             TDVLDFYLEKNVNVVIPMEGKFSYYTDWVEENASLGGKQMWETFLVKELPGPLEEKLNTD
Sbjct  121   TDVLDFYLEKNVNVVIPMEGKFSYYTDWVEENASLGGKQMWETFLVKELPGPLEEKLNTD
180

Query  575  GQRAIAGMSMSATTSLLFPQHFPGFYDAAASFSGCAATSSLLPWEYLKLTLDRGNATPEQ
754

             GQRAIAGMSMSATTSLLFPQHFPGFYDAAASFSGCAATSSLLPWEYLKLTLDRGNATPEQ
Sbjct  181   GQRAIAGMSMSATTSLLFPQHFPGFYDAAASFSGCAATSSLLPWEYLKLTLDRGNATPEQ
240

Query  755  MWG  763
             MWG
Sbjct  241  MWG  243
```

Rev

Range 1: 30 to 365GenPeptGraphics Next Match Previous Match First Match

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
609 bits(1570)	0.0()	Compositional matrix adjust.	299/337(89%)	303/337(89%)	1/337(0%)	+3

Features:

```
Query  189  AEVTPADVAGDNCTIHHLRIVLLQMKPLHSLARHTSTQQTSASKKCGHTPLSMDRNVPLV
368

             AEVTPADVAGD                        H +       +   ++P SMDRNVPLV
Sbjct  30    AEVTPADVAGDTALSTISDSAPADEASAPRWRAHVNAADERVKEMWAYSP-SMDRNVPLV
88

Query  369  VITADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPMEGKFSYYTDW
548

             VITADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPMEGKFSYYTDW
Sbjct  89    VITADESAGPRPVIYLLNGGDGGEGAANWVMQTDVLDFYLEKNVNVVIPMEGKFSYYTDW
148

Query  549  VEENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLFPQHFPGFYDA
728

             VEENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLFPQHFPGFYDA
Sbjct  149   VEENASLGGKQMWETFLVKELPGPLEEKLNTDGQRAIAGMSMSATTSLLFPQHFPGFYDA
208

Query  729  AASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALINSDKLRGTELY
908

             AASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALINSDKLRGTELY
Sbjct  209   AASFSGCAATSSLLPWEYLKLTLDRGNATPEQMWGPRGGEYNIYNDALINSDKLRGTELY
268

Query  909  VSNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTHDLKAKLDSAG
1088

             VSNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTHDLKAKLDSAG
```

```
Sbjct  269   VSNASGLAGEWESVDSPRFEGLNQQVQSIAMAETVVTGGIIEAATNKCTHDLKAKLDSAG
328

Query  1089  IPADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA  1199
             IPADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA
Sbjct  329   IPADWNLRPTGTHSWGWWQDDLRGSWTTFARAFELEA  365
```

Figure L1 Sequence trace of VapE optimised for expression in *E. coli* BL21 (DE3)

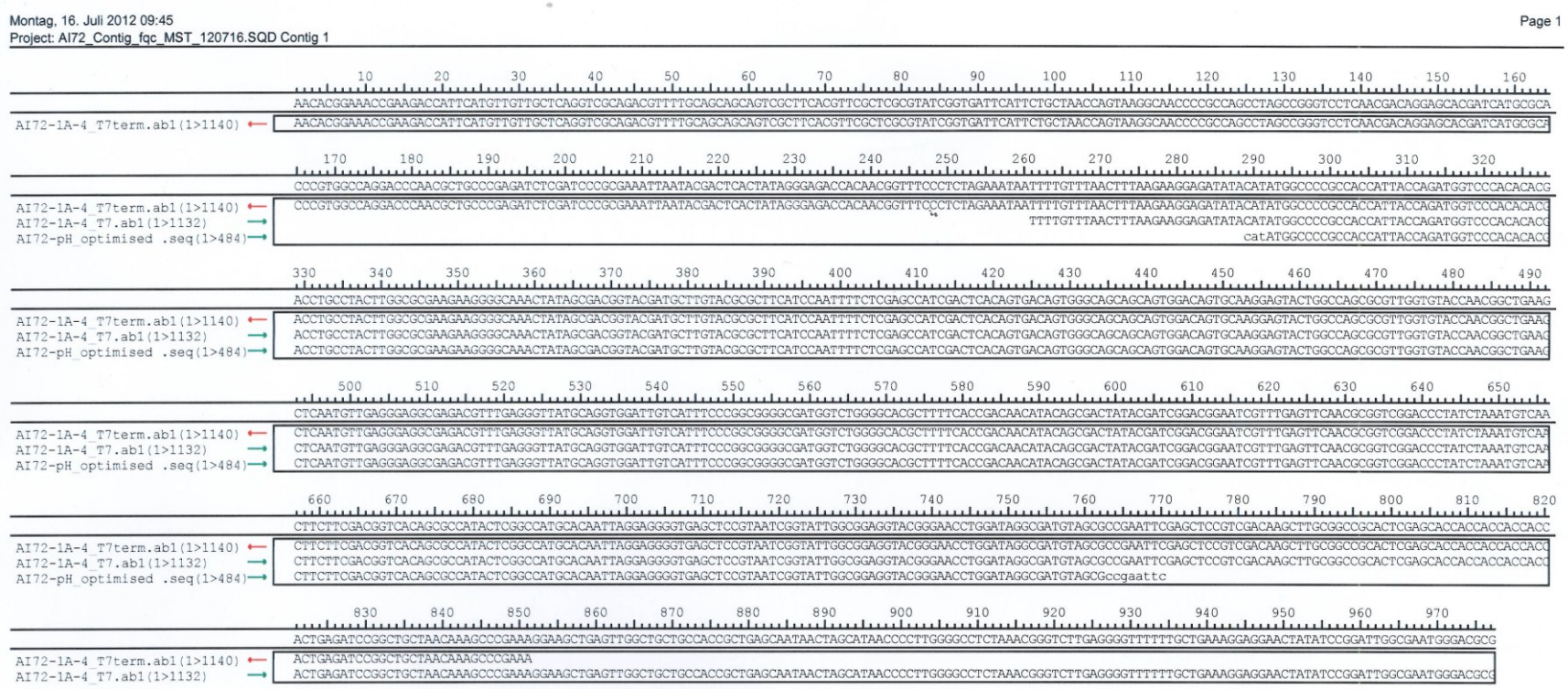


Figure L2 Sequence trace of VapH optimised for expression in *E. coli* BL21 (DE3)

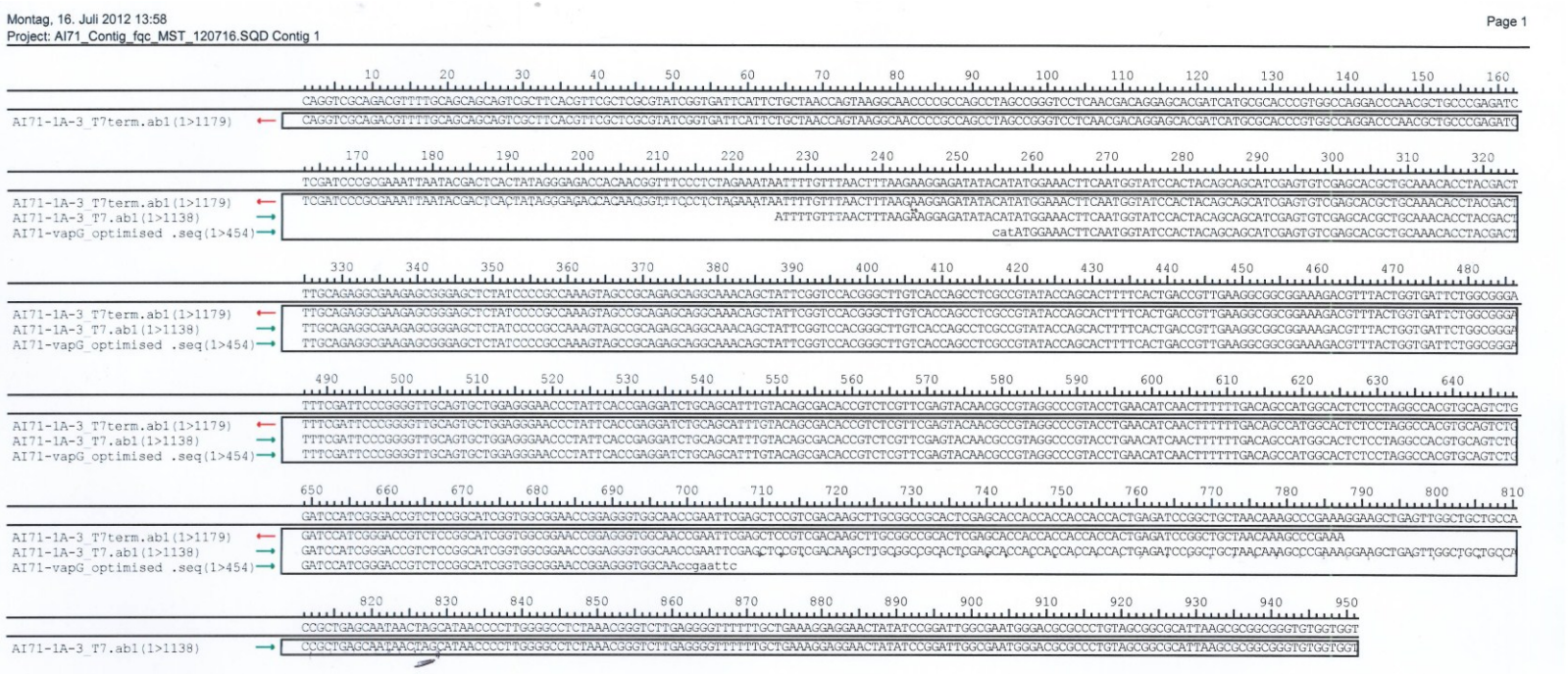


Figure L3 Sequence trace of VapG optimised for expression in *E. coli* BL21 (DE3)